

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2005, 00:56:52 ; Search time 405 Seconds

(without alignments)  
6856.651 Million cell updates/sec

Title: US-10-725-829-1

Perfect score: 529  
Sequence: 1 cataatggctgttcacaag.....ctgtaatggactttattg 529

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_23Sep04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	529	100.0	529	12	ADP02854 Solanum a
2	288.4	54.5	632	12	ADP02856 Solanum a
3	259.2	49.0	684	10	ADG05323 Tin2 stru
4	258	48.8	1241	2	AAQ02053 wound-ind
5	254.2	48.1	584	3	AZA49862 Potato pr
6	178.8	33.8	1360	2	AAQ68729 Full leng
7	149	28.2	1104	2	AAQ68728 Nucleotid
8	149	28.2	1104	6	AD156863 Plant def
9	35.8	6.8	2000	8	ADA71938 Rice gene
10	35.6	6.7	48727	4	AAK67375 Human imm
11	35.6	6.7	108359	9	ADA13316 Human fri
12	35.2	6.7	6134	4	AAS46563 Tumour su
13	35.2	6.7	6650	6	ABL32638 Human imm
14	35.2	6.7	10433	6	ABL232379 Human imm
15	35	6.6	9964	6	ABL32098 Human imm
16	34.8	6.6	2591	4	ABL08938 Human imm
17	34.8	6.6	6866	6	ABL32666 Human imm
18	34.8	6.6	6866	6	ABL49319 Human pol
19	34.6	6.5	346	6	ABN94362 Gene #860
20	34.4	6.5	2905	12	ADJ75932 Marker ge
21	34.4	6.5	63155	10	ADC85996 Human GPC

ALIGNMENTS

RESULT 1  
ADP02854  
ID ADP02854 standard; DNA; 529 BP.  
XX  
AC ADP02854;  
XX  
DT 09-SEP-2004 (first entry)  
XX  
DE Solanum americanum proteinase inhibitor II gene SapIN2a.

ds; gene; proteinase inhibitor II; transformed plant; resistance; insect;  
pest; pathogen; programmed cell death; senescence.  
XX  
OS Solanum americanum.

XX Key Location/Qualifiers  
FH 5. .451  
CDS /\*tag= a  
FT /product= "SapIN2a protein"

XX WO2004050873-A1.  
XX 17-JUN-2004.

XX 01-DEC-2003; 2003WO-CN001020.

XX 29-NOV-2002; 2002US-0429992P.

XX (UYHK-) UNIV HONG KONG.

XX Chye M, Xu Z, Sin S;

XX WPI; 2004-450731/42.

XX P-PSDB; ADP02855.

XX New proteinase inhibitor, SapIN2a or SapIN2b, useful in producing transformed plants having enhanced resistance to insects, pest or pathogens and in which programmed cell death or senescence is inhibited.  
XX Claim 1; SEQ ID NO 1; 90pp; English.

XX The invention relates to an isolated proteinase inhibitor II nucleic acid molecule (I). The proteinase inhibitor II nucleic acid molecule and the encoded polypeptide, methods are useful in producing transformed plants

Continuation (20 o  
Adf13122 Hypermeth  
Adf13115 Hypermeth  
Adf137268 Hypermeth  
Adf137261 Hypermeth  
Aax13696 Enterococ  
Abe99491 Enterococ  
Abl04388 Drosophil  
Aal04789 Human rep  
Abl197684 Human tes  
Abk85261 Human gen  
Adf13990 Human pro  
Aai58381 Human pol  
Adq98590 DNA encod  
Aab48350 Novel hum  
Aal50282 Human nuc  
Adc76022 DNA homol  
Adk55393 Plant DNA  
Aah53990 S. epider  
Aak81662 Human imm  
Aak82628 Human imm  
Aak81663 Human imm  
Aak70161 Human imm  
Aak81666 Human imm

CC	having enhanced resistance to insects, pest or pathogens and in which
CC	programmed cell death or senescence is inhibited. This sequence
CC	corresponds to the American black nightshade <i>SapIN2a</i> gene.
XX	
SQ	Sequence 529 BP; 155 A; 97 C; 112 G; 165 T; 0 U; 0 Other;
Query Match	
Best Local Similarity 100.0%; Score 529; DB 12; Length 529;	
Matches 529; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 CATAATGGCTGTTCCACAAAGTTAGCTTCCTTGGCTTACTTGTCTTGGATGGATGTT 60
Db	1 CATAATGGCTGTTCCACAAAGTTAGCTTCCTTGGCTTACTTGTCTTGGATGGATGTT 60
Qy	61 TCTACTTTGCCAAAACATGTTGATGCCAAGGCTTGTACTAGAGAAATGCTGCTATTTAGCTA 120
Db	61 TCTACTTTGCCAAAACATGTTGATGCCAAGGCTTGTACTAGAGAAATGCTGCTATTTAGCTA 120
Qy	121 TGGCATATGCCCAGTTCAGAAAGGAAGTCCCAAAAACCTATATGCAACCAATTTGTTGCTC 180
Db	121 TGGCATATGCCCAGTTCAGAAAGGAAGTCCCAAAAACCTATATGCAACCAATTTGTTGCTC 180
Qy	181 AGGCTATATAGGGTTGCAACTATTACAGTGTCTAAAGGAGATTTGATTTGTGAAGGAGATC 240
Db	181 AGGCTATATAGGGTTGCAACTATTACAGTGTCTAAAGGAGATTTGATTTGTGAAGGAGATC 240
Qy	241 TGACCCCTAGAAACCCAAAAGATTTGACCTTCGAATGTGATACACAGATTGCTTATTCAAA 300
Db	241 TGACCCCTAGAAACCCAAAAGATTTGACCTTCGAATGTGATACACAGATTGCTTATTCAAA 300
Qy	301 ATGTCCTCGTTTCAGAAAGAAAGATGATAATTTAAACCCCACTGGATGCCACCACTTGTGCGAC 360
Db	301 ATGTCCTCGTTTCAGAAAGAAAGATGATAATTTAAACCCCACTGGATGCCACCACTTGTGCGAC 360
Qy	361 GGGCTATCAGGGTTGCTACTATTTCGATCAAGATGGTGATTTTGTCTGTGAAGGAGAG 420
Db	361 GGGCTATCAGGGTTGCTACTATTTCGATCAAGATGGTGATTTTGTCTGTGAAGGAGAG 420
Qy	421 TCCTGAAACCCCAAGACCACTGCCTTATTTCTAATCAATCATATGTTGTTATCTATCAAAAAA 480
Db	421 TCCTGAAACCCCAAGACCACTGCCTTATTTCTAATCAATCATATGTTGTTATCTATCAAAAAA 480
Qy	481 AAATATGATGATCATGATATATGCTGGTTACTGTAATGTTGACTTTATTTG 529
Db	481 AAATATGATGATGATATATGCTGGTTACTGTAATGTTGACTTTATTTG 529
RESULT 2	
ADP02856	
ID	ADP02856 standard; DNA; 692 BP.
XX	
AC	ADP02856;
XX	
DT	09-SEP-2004 (first entry)
XX	
DE	Solanum americanum proteinase inhibitor II gene SapIN2B.
XX	
KW	ds; gene; proteinase inhibitor II; transformed plant; resistance; insect;
KW	pest; pathogen; programmed cell death; senescence.
OS	Solanum americanum.
XX	
PH	Key Location/Qualifiers
FT	CDS 33..491
FT	/*tag= a
FT	/product= "SapIN2b protein"
XX	
PN	WO2004050873-A1.
XX	
PD	17-JUN-2004.
XX	
PF	01-DEC-2003; 2003WO-CN001020.
XX	

[illegible]

Qy	131	CCACGTTCAGAAAGTCCCCAAAAA	CCTATATGCA	CCAAATGTTGCTCAGGCTAT	AAG	190
Db	558	CCACGTTCAGAAAGTCCGGAAATCCGATATGCACCAACTGTTGTGCAGGTTAT	AAA	617		
Qy	191	GGTTGCAACTATTACAGTGCTAAAGAGATTTGATTTGTGAAGGAGAACTCGACCCCTAG	A	250		
Db	618	GGTTGCAATTATTATGTGCAAAATGGGGCTTTCATTTGTGAAGGACAATCTCGACCCAAAA	A	677		
Qy	251	AACCCAAAAGATGTGACCTTCGAATGTGATACACAGATTGCTATTTCAAAAATGTCCTCGT	A	310		
Db	678	AAACCAAAAGCATGCCCCCTAAATTCGATCCACATATTGCCTACTCAAAGTGTCCCCTG	A	737		
Qy	311	TCAGAAGAAAGATGATTAATTAACCCCACTGGATGCACCACTGTTGTGACGGGCTATCAG	A	370		
Db	738	TCAGAAGGAAAATCGCTAAATTTATCCCAACCGGATGTACCACTGTGTGACAGGGTACAG	A	797		
Qy	371	GGTTGCTACTATTTCGATCAAGATGGTGATTTTGTCTGTGAAGGAGAGAGTCTCGAACCC	A	430		
Db	798	GGTTGCTACTATTTCGGTTAAAAATGCGCAAGTTTGTATGTGAAGGAGAGATGTATGAGCCC	A	857		
Qy	431	AAGACCACTGCTTA	444			
Db	858	AAGGCAATATGTA	871			

RESULT 5	
AAZ49862	
ID	AAZ49862 standard; DNA; 584 BP.
XX	
XX	
AC	AAZ49862;
XX	
DT	25-APR-2000 (first entry)
XX	
XX	Potato proteinase inhibitor-II DNA.
DE	
XX	
KW	Potato proteinase inhibitor-II; PPI-II; strepavidin; worm; insect;
KW	plant-nnoxious protein; pest resistance; moth; insect; weevil; grub;
KW	beetle; fly; thrip; locust; cricket; borer; mite; looper; insecticidal;
KW	ss.
XX	
XX	
OS	Solanum tuberosum.

Christeller JT, Sutherland PW, Murray C, Markwick NP, Philip BA;  
Malone LA, Burgess EPU;  
WPI; 2000-171244/15.

New chimeric polypeptide and composition comprising the polypeptide  
useful for conferring pest resistance on plants.

Example 3; Fig 4; 11lpp; English.

The present sequence encodes potato proteinase inhibitor-II (PPI-II).  
This is used in the preparation of a binary vector designed to express a  
chimeric polypeptide comprising streptavidin mature peptide, a plant -  
noxious protein, fused to the PPI-II signal peptide. The binary vector is  
targeted to the vacuole by PPI-II signal sequence. Transformation of  
plant genome with the vector can produce pest resistance in plants, plant  
derived products and stored harvest material. Pests that can be  
controlled include, cotton bollworm, tropical army-worm, European corn -  
borer or red mite, tobacco horn worm, loopers, rice stem borer, porina,  
cutworms, diamondback moth, potato tuber moth, codling moth, Indian meal  
moth, gypsy moth, argentine stem weevil, clover root weevil, grass -  
grubs, corn rootworm, rice and wheat weevils, mealworms, flour beetles,  
black field cricket, locusts, sawflies, Western flower thrips, Hessian  
flies or two-spotted mite

Sequence 584 BP; 184 A; 94 C; 116 G; 190 T; 0 U; 0 Other;

Query Match      48.1%; Score 254.2; DB 3; Length 584;  
Best Local Similarity 79.4%; Pred.No. 5.7e-66;  
Matches 301; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

Qy      66    TTGCGAAACATGTTGATGCCAAGGCTTGTA CTAGAGAATGTGTCATTTTAGCTATGGCA 125  
         ||||| |  
Db      193    TGGTGGAACATGTTGATGCGAGATCTGTA AAGAATGTGTAATCTTGGGTTGGGA 252  
         ||||| |  
Qy      126    TATGCCCA CGTTCAGAAAGGAAGTCCCAA AAA CCTATATGACC CAATGTGTGCTCAGGCT 185  
         ||||| |  
Db      253    TATGCCCAC GTTCAGAAAGGAAGTCCGAAA AATCCCATATGCATCAATGTGTGCTCAGGCT 312  
         ||||| |  
Qy      186    ATAAGGGTGC CACTATTACAGTGCTAA AGGAGATTTTGATTTGTGAAGGAGAATCTGACC 245  
         ||||| |  
Db      313    ATAAGGGTGT GAATATTATAGTGT TTCGGGAGATTTATTTCGAAGGAGAATCTGACC 372  
         ||||| |  
Qy      246    CTAGAAACCC AAAGATTTGTAACCTTCG AATGTGATACACAGATGTCTTATTCAAAATGTC 305  
         ||||| |  
Db      373    TAAAAACCC AAAGCTTGCCCCCTAA ATTTGTGATACAAATATTTGCCTATTCAGATGCC 432  
         ||||| |  
Qy      306    CTCGTTT CAGAAAGGAAGATGATAA TAA ACCCATCGGATGCCACA CTGTTGTGCAAGGCT 365  
         ||||| |  
Db      433    CCCATTTCAG AAGGAAAAATCGCTAA TTTATTC CCACCGGATGTACCACATGTTGCACAGGT 492  
         ||||| |  
Qy      366    ATCAGGGTTC TACTATTTCGATCAAGAT GGTGATTTTGCTGTGAAGGAGAGAGTCCGTG 425  
         ||||| |  
Db      493    ACAAGGGTGT CTAATTTCGTATAAAAT TGCCAAGTTTGTATGCGAAGGAGAGAGTGATG 552  
         ||||| |  
Qy      426    AACCCAAGAC CACTGCTTA 444  
         ||||| |  
Db      553    AACCCAAGCA ATATGTA 571  
         ||||| |

RESULT 6  
AAQ68729  
ID    AAQ68729 standard; DNA; 1360 BP.  
XX    AC    AAQ68729;  
XX    AC     
DT    25-MAR-2003 (revised)  
DT    02-MAR-1995 (first entry)  
XX    Full length sequence of PI precursor.  
DE    Type II serine proteinase inhibitor precursor; PI; tobacco;  
KW    transgenic plant; anti-pathogen; anti-predator; ss.  
XX





181 TGTACCTTAAACTGTGATCCAGAAATTCCTATGGAGTTTCCCGCGTTTCAGAGAAAG 240  
323 ATGATAATTAACCCACTGGATGCACCACATTTGTCACCGGCTATCAGGGTTCGTACTAT 382  
241 AAG-----AATGATCGGATATGACCACTGTTGCGCAGGACGAAAGGTTTGAAGTAC 294  
383 TTCGATCAAGATGATGATTTGTCGTGAAGGAGAGAGTCTCTGAACCCCAAGACCACTGCT 442  
295 TTCAGTATGATGGAACCTTTGTTGTAAGGAGAGAGTCTGATCCTAGAAATCCAAAGGCT 354  
443 TAT 445  
355 TGT 357

RESULT 8  
ADI56863  
ID ADI56863 standard; DNA; 1104 BP.  
XX AC ADI56863;  
XX XX  
XX DE 22-APR-2004 (first entry)  
XX DE Plant defensin NaPI mature domain cDNA.  
XX ds; gene; insecticide; antifungal; virucide; antibacterial;  
XX flower development; plant pest resistance; floral defensin precursor;  
XX transgenic plan; insect resistance.  
XX Unidentified.  
XX XX  
XX PN WO200263011-A1.  
XX PD 15-AUG-2002.  
XX XX  
XX PF 08-FEB-2002; 2002WO-AU000123.  
XX XX  
XX PR 08-FEB-2001; 2001US-0267271P.  
XX XX  
XX PA (HEXI-) HEXIMA LTD.  
XX XX  
XX Anderson MA, Lay FT, Heath RL;  
XX PI  
XX WPI; 2002-657538/70.  
XX DR P-PSDB; ADI56864.  
XX XX  
XX New isolated plant floral defensin-like polypeptides and polynucleotides,  
XX for generating transgenic plants having resistance or at least reduced  
XX sensitivity to plant pests including insects, microorganisms, fungi  
XX and/or viruses.  
XX  
XX Disclosure; SEQ ID NO 56; 164pp; English.  
XX XX  
XX The invention relates to an isolated polypeptide comprising, in its  
XX precursor form, an N-terminal signal domain, a mature domain and an  
XX acidic C-terminal domain, where the polypeptide is produced during flower  
XX development and its mature domain has activity against one or more plant  
XX pests. Plant floral defensin-like polypeptides and polynucleotides are  
XX useful in generating transgenic plants having resistance or at least  
XX reduced sensitivity to plant pests including insects, microorganisms,  
XX fungi and/or viruses. They are also useful in generating recombinant  
XX defensin-like molecules for use in the topical application of  
XX compositions to prevent or retard pest-infestation of plants. The floral-  
XX and seed-derived defensins are useful in the generation of insect  
XX resistance in plants. This sequence represents the cDNA encoding the  
XX mature domain from a plant defensin protein.  
XX  
XX Sequence 1104 BP; 337 A; 176 C; 295 G; 296 T; 0 U; 0 Other;  
XX  
XX Query Match 28.2%; Score 149; DB 6; Length 1104;  
XX Best Local Similarity 67.2%; Pred. No. 3.3e-34;  
XX Matches 244; Conservative 0; Mismatches 110; Indels 9; Gaps 2;

QY 86 AAGGCTTGTACTAGAGAATGTGTC---ATTTTAGCTATGGCATATGCCACCGTTCAGAA 142  
DB 1 AAGGCTTGTACTTAAACTGTGATCCAGAAATTTGCCCTATGGAGTTTCCCGCGTTTCAGAA 60  
QY 143 GGAAGTCCCAAAACCTATATATGCACCAATTTGTCAGGCTATTAAGGTTTGCACACTAT 202  
DB 61 GAAAAGAAAGATGATCGGATATGCACCAACTGTTGGCAGGACGAAAGGTTTGAAGTAC 120  
QY 203 TACAGTGTCTAAAGGAGATTTGATTTCGTGAAGGAGAAATCTGACCTAGAAAACCCAAAAGAT 262  
DB 121 TTCAAGTATGATGGAACCTTTGTTGTAAGGAGAGTCTGATCCTAGAAATCCAAAGGCT 180  
QY 263 TGTACCTTCGAATGTGATACACAGATTGCTTATTTCAAAATGTCTCTGTTTCAGAGGAAAG 322  
DB 181 TGTACCTTAAACTGTGATCCAGAAATTTGCTATGGAGTTTCCCGCGTTTCAGAGGAAAG 240  
QY 323 ATGATAATTAACCCACTGGATGCACCACTTGTGTCACCGGCTATCAGGGTTCGTACTAT 382  
DB 241 AAG-----AATGATCGGATATGCACCAACTGTTGGCAGGACGAAAGGTTTGAAGTAC 294  
QY 383 TTCGATCAAGATGATGATTTGTCGTGAAGGAGAGAGTCTCTGAACCCCAAGACCACTGCT 442  
DB 295 TTCAGTATGATGGAACCTTTGTTGTAAGGAGAGTCTGATCCTAGAAATCCAAAGGCT 354  
QY 443 TAT 445  
DB 355 TGT 357

RESULT 9  
ADA71938/c  
ID ADA71938 standard; DNA; 2000 BP.  
XX AC ADA71938;  
XX XX  
XX DT 20-NOV-2003 (first entry)  
XX DE Rice gene, SEQ ID 5263.  
XX DE Plant; bacterial infection; fungal infection; viral infection; rice;  
XX KW gene; ds.  
XX OS Oryza sativa.  
XX XX  
XX PN WO2003000898-A1.  
XX XX  
XX PD 03-JAN-2003.  
XX XX  
XX PF 22-JUN-2001; 2001WO-IB001105.  
XX XX  
XX PR 22-JUN-2001; 2001WO-IB001105.  
XX XX  
XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX XX  
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
XX Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
XX WPI; 2003-175290/17.  
XX  
XX Identifying at least one gene involved in plant resistance or response to  
XX pathogenic infection for conferring resistance or tolerance to a plant to  
XX bacterial, fungal or viral infection by determining or detecting plant  
XX gene expression.  
XX  
XX Claim 27; SEQ ID NO 5263; 899pp; English.  
XX  
XX The present invention relates to a method (M1) for identifying genes  
XX involved in plant resistance or response to pathogenic infection. M1  
XX comprises identifying a gene whose expression is significantly altered in  
XX the incompatible interaction of plant gene expression relative to  
XX expression of the gene in an uninfected plant, in a mutant plant that  
XX does not express a gene associated with response to pathogenic infection,



```
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246479P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 03-DEC-2000; 2000US-0254097P.
PR 03-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
DR
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and metastasis.
PT
XX Disclosure; SEQ ID NO 22187; 307lpp + Sequence Listing; English.
PS
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
```

```
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 48727 BP; 16775 A; 9367 C; 9223 G; 13362 T; 0 U; 0 Other;

Query Match      6.7%; Score 35.6; DB 4; Length 48727;
Best Local Similarity 53.6%; Pred. No. 21;
Matches 74; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 246 CTAGAAACCCAAAGATTGTCCTTCGAATGTGATACACAGATTGCTTATTCAAAATGTC 305
Db 3174 CTTGAGACCTTAAGAGATTGCTTCTTGGAAATAAATACAGGTTTATAGATGAC 3115
QY 306 CTCGTTTCAGAGGAAAGATGATAATTAACCCACTGCACCACTTGTTCACCGGCT 365
Db 3114 CTAATTAAAGTATACAGGTTATGAGAGAGGCTTGTGAAGCGCTTGTATGATGATGCTT 3055
QY 366 ATCAGGGTTGCTACTATT 383
Db 3054 TTAAGTTTACACCTTTT 3037

RESULT 11
ADA13316/C
ID ADA13316 standard; DNA; 108359 BP.
XX
AC ADA13316;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human fringe-like secreted protein gene, SEQ ID NO:3.
XX
KW Human; secreted protein; glycosyltransferase family; fringe protein;
KW testis; hepatocellular carcinoma; placenta; germinal centre B cell;
KW brain; chromosome 13; drug screening; cancer; cytostatic; gene therapy;
KW biochip; transgenic animal; gene; ds.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT variation replace(2220,A)
FT /*tag= a
FT /note= "Single nucleotide polymorphism (SNP)"
FT CDS 2999..103512
FT /*tag= c
FT /partial
FT /product= "Human fringe-like secreted protein"
FT /note= "Encodes residues 53-499 of human fringe-like
FT secreted protein; no start codon given; contains 11
FT introns"
FT exon 2999..3112
FT /*tag= b
FT /transl_except= (pos:2999..3001;aa:Ile)
FT /transl_except= (pos:3101..3103;aa:His)
FT intron 3113..320842
FT /*tag= d
FT variation replace(3639,G)
FT /*tag= e
FT variation replace(3722,A)
FT /*tag= f
FT variation replace(4673,G)
FT /*tag= g
FT variation replace(6235..6237,Gr)
FT /*tag= h
FT variation replace(11368,T)
FT /*tag= i
FT variation replace(12969,T)
```



```
FT intron 48453. .50526
FT /*tag= bj
FT variation replace(49219,C)
FT /*tag= bk
FT /note= "Single nucleotide polymorphism (SNP)"
FT replace(49259,A)

Query Match 6.7%; Score 35.6; DB 9; Length 108359;
Best Local Similarity 45.9%; Pred. No. 30;
Matches 122; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

QY 235 AGAATCGACCTAGAACCCAAAGATTCTGATGATGATACACAGATTGCTTA 294
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 850 ACAATATAAACAATAATATATAATGATATATACCAACATATACACATAGCATG 791
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 295 TTCAAAATGCTCGTTCAGAGGAAGATGATAATTAACCCACCTGGATGCCACTTG 354
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 790 TGAAGGTATAAAGATTACGTGAGTATGATAAATACCAATCAGGATAATGGTTAAAGG 731
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 355 TTGCACGGCTATCAGGGTTGCTACTATTTTCGATCAAGATGGTGATTTCTGCTGAAGG 414
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 730 TTTCAAGTGTGTTAATGGTATTTAATTTCTTAATCTAGGTGATATTTTCTTAACTCT 671
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 415 AGAGAGTCTGGAACCCAGACCACTGCTTATTTCTTAATCAATCATATGTTGTTATCTATC 474
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 670 AATCAACATATAATACATGAATGTTCAATATATATTTCTTATTAACCTCTTATATGTT 611
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 475 AAAAAAATATGTCATGATATA 500
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 610 CAAATATTCTTAATACTTTTTTTA 585
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
AAS46563/C
ID AAS46563 standard; DNA; 6134 BP.
XX AC AAS46563;
XX DT 18-DEC-2001 (first entry)
XX DE Tumour suppressor gene derived chemically modified sequence #285.
XX KW Human; tumour suppressor gene; oncogene; antitumour; cytostatic; cancer;
XX KW tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
XX KW cytosine methylation; ds.
XX OS Homo sapiens.
XX PN WO200168912-A2.
XX PD 20-SEP-2001.
XX PF 15-MAR-2001; 2001WO-EP002955.
XX PR 15-MAR-2000; 2000DE-01013847.
XX PR 06-APR-2000; 2000DE-01019058.
XX PR 07-APR-2000; 2000DE-01019173.
XX PR 30-JUN-2000; 2000DE-01032529.
XX PR 01-SEP-2000; 2000DE-01043826.
XX PA (BPIG-) EPIGENOMICS AG.
XX PI Olek A, Piepenbrock C, Berlin K;
XX WP; 2001-602752/68.
XX Fragments of chemically modified genes associated with tumor suppressor
XX PT genes and oncogenes, useful in designing primers and probes for analyzing
XX PT diseases associated with cytosine methylation state e.g. cancer.
XX PS Claim 1; SEQ ID NO 285; 27pp; English.
XX CC The invention relates to a nucleic acid comprising a sequence of 18
```

```
CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
CC bisulphite, of genes associated with tumour suppression and oncogenes
CC having a sequence taken from 536 (actually 533 since numbers 408, 458 and
CC 500 are missing from the sequence listing) sequences (Ss) and sequences
CC complementary to (Ss). The nucleic acid may be a peptide nucleic acid-
CC oligomer (PNA) of at least 9 nucleotides and may form part of a set of
CC probes for detecting the cytosine methylation state and/or single
CC nucleotide polymorphisms and also to be used in an array for analysing
CC diseases associated with CpG dinucleotides e.g. cancers and tumours. The
CC probes can also be used in a method for ascertaining genetic and/or
CC epigenetic parameters for the diagnosis and/or therapy of existing
CC diseases or the predisposition to specific diseases, by analysing
CC cytosine methylations. The parameters may be compared to another set of
CC genetic and/or epigenetic parameters, the differences serving as basis
CC for diagnosis and/or prognosis events which are disadvantageous to
CC patients. The present sequence is one of the 533 genomic sequences
CC derived from tumour suppressor genes and oncogenes. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 6134 BP; 1576 A; 76 C; 1372 G; 3110 T; 0 U; 0 Other;

Query Match 6.7%; Score 35.2; DB 4; Length 6134;
Best Local Similarity 44.9%; Pred. No. 12;
Matches 133; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

QY 220 TTTGATTTGTGAGGAGAAATCTGACCCCTAGAACCCCAAGATTGTACCTTCGAATGTGA 279
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 4083 TTCGAAATTTTACAAAATAAATAACATTAATAAATAAATAAATAAATAAATAAATAA 4024
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 280 TACACAGATTGCTTATTCAAAATGTCCTCGTTCCAGAGGAAGATGATATTAACCCAC 339
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 4023 ATATAAAATTACATCTACATATAACCTCAATACAAAAAATAAATAAATAAATAAATAA 3964
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 340 TGGATGCACCACTTGTTCACGGGCTATCAGGGTTCCTACTATTTTCGATCAAGATGTGA 399
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3963 AAAAAATATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3904
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 400 TTTTGTCTGTGAGGAGAGAGTCTCGAACCCAGACCACTGCTTATTTCTTAATCAATCAT 459
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3903 CTTTCTCTATTTTTCATTAACCTTCAACACAAAAACCATACATATTTATATAAAAAA 3844
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 460 ATGTTGTTATCTATCAAAAAAATAATATGATGATGATATGCTGTTACTGTAA 515
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3843 AATAATTTATTTTAACTAACATAAATAAATAAATAAATAAATAAATAAATAAATAA 3788
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
ABL32638
ID ABL32638 standard; DNA; 6650 BP.
XX AC ABL32638;
XX DT 26-MAR-2002 (first entry)
XX DE Human immune system associated gene SEQ ID NO: 611.
XX KW Human; immune system disease; cytosine methylation; antiasthmatic;
XX KW antiarteriosclerotic; antianaemic; cytosstatic; nootropic;
XX KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
XX KW antiinflammatory; antiarthritic; antidiabetic; antipsoriatic;
XX KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
XX KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
XX KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
XX KW ds.
XX OS Homo sapiens.
XX PN WO200200928-A2.
XX PD 03-JAN-2002.
XX
```



XX Nucleic acid comprising fragment of chemically modified gene, useful for  
PT diagnosis and treatment of diseases associated with abnormal cytosine  
PT methylation.  
XX  
PS Claim 1; SEQ ID NO 71; 32pp + Sequence Listing; German.  
XX  
CC The present invention provides a number of human immune system associated  
CC genes which are modified by the methylation of cytosines. The sequences  
CC can be used in the diagnosis and treatment of immune system disorders,  
CC including eye diseases such as retinopathy, neovascular glaucoma and  
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
CC diseases. The present sequence is a gene of the invention  
XX  
SQ Sequence 9964 BP; 2847 A; 42 C; 1995 G; 5080 T; 0 U; 0 Other;  
  
Query Match 6.6%; Score 35; DB 6; Length 9964;  
Best Local Similarity 66.7%; Pred. No. 17;  
Matches 50; Conservative 0; Mismatches 25; Indels 0; Gaps 0;  
  
Qy 455 ATCATATGTTGTTATCTATCAAAAAAATATGTCATGATATATGCTGTTACTGTA 514  
Db 3825 ATTTAGGATGTTATTTTAAAAAATGTTTTATAAATGTTAAAGATGTTTATTGTA 3884  
  
Qy 515 ATGTGGACTTTATTG 529  
Db 3885 ATATTAAATTATATG 3899

Search completed: January 22, 2005, 05:47:32  
Job time : 414 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2005, 01:01:52 ; Search time 2760 Seconds

(without alignments)  
9063.859 Million cell updates/sec

Title: US-10-725-829-1

Perfect score: 529

Sequence: 1 cataatggctgtcacaaag.....ctgtaatgtggactttatttg 529

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_hgt.\*

3: gb\_in.\*

4: gb\_cm.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	529	100.0	529	8	AF174381 Solanum a
2	288	54.4	687	8	AY422686 Solanum n
3	288	54.4	766	8	LECSVI57G X94946 L.esculentu
4	279.6	52.9	838	8	STPIN2W X99095 S.tuberosum
5	265.4	50.2	660	8	TOMWPII K03291 Tomato leaf
6	265.4	50.2	1776	8	BT013250 Lycopersi
7	263.6	49.8	1695	8	STPRIN2G Z12753 S.tuberosum
8	260.6	49.3	2068	8	STU45450 Solanum tub
9	259.2	49.0	559	8	AB110700 Lycopersi
10	259.2	49.0	684	8	AY007240 Lycopersi
11	259.2	49.0	1670	8	AY129402 Lycopersi
12	258	48.8	1241	8	POTI1KA M15186 S.tuberosum
13	257.6	48.7	482	8	AY247794 Solanum p
14	256.8	48.5	512	8	STPIIIR1 X03778 Potato (Sol
15	256.8	48.5	666	8	POTPPINHB L37519 Solanum tub
16	256.4	48.5	1274	8	NTPROTINH Z29537 N.tabacum (
17	255.8	48.4	2330	8	STPINI1 Z13992 S.tuberosum
18	254.2	48.1	584	8	STPIN2 X78275 S.tuberosum
19	251.8	47.6	580	8	AY517498 Solanum p

20	250.2	47.3	1914	8	STPI2G X04118 Potato gene
21	249	47.1	554	8	STPIIIR2 X03779 Potato (Sol
22	227.4	43.0	532	8	AF209709 Solanum a
23	219.4	41.5	1573	8	STPRINPSG Z12824 S.tuberosum
24	178.8	33.8	1342	8	NAU08219 U08219 Nicotiana a
25	178.8	33.8	1360	6	E54395 Transgenic
26	178.8	33.8	1360	6	AR224422 Sequence
27	178.8	33.8	1360	6	AR230263 Sequence
28	175.6	33.2	946	8	AF105340 Nicotiana
29	164.4	31.1	1546	8	AF542547 Nicotiana
30	158	29.9	1795	8	AF205851 Nicotiana
31	149.6	28.3	764	8	AF221097 Capsicum
32	149	28.2	1104	6	E54394 Transgenic
33	149	28.2	1104	6	AR224421 Sequence
34	149	28.2	1104	6	AR230262 Sequence
35	148.8	28.1	1329	8	AF205852 Nicotiana
36	143.8	27.2	836	8	AF039398 Capsicum
37	132.8	25.1	1414	8	AY426751 Nicotiana
38	114.2	21.6	573	8	AF242734 Capsicum
39	105.8	20.0	562	8	BT013127 Lycopersi
40	93	17.6	841	8	L21194 Lycopersico
41	87.8	16.6	4272	8	TOMARPI L21128 Lycopersico
42	66.2	12.5	175	8	AY184823 Nicotiana
43	42.8	8.1	1379	5	CR733321 Nicotiana
44	42.2	8.0	536	11	G58878 Gallus ga
45	42	7.9	235502	2	AC103222 Rattus no

#### ALIGNMENTS

RESULT 1	AF174381	529 bp	mRNA	linear	PLN 15-JAN-2002
LOCUS	Solanum americanum	proteinase inhibitor IIA (PIN2a)	mrna	complete	
DEFINITION	cda.				
ACCESSION	AF174381				
VERSION	AF174381.1	GI:17221674			
KEYWORDS	Solanum americanum				
SOURCE	Solanum americanum				
ORGANISM	Solanum americanum				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.				
AUTHORS	Xu, Z.F., Qi, W.Q., Ouyang, X.Z., Yeung, B. and Chye, M.L.				
TITLE	A proteinase inhibitor II of Solanum americanum is expressed in phloem				
JOURNAL	Plant Mol. Biol.	47 (6), 727-738 (2001)			
MEDLINE	21643209				
PUBMED	11785934				
REFERENCE	2 (bases 1 to 529)				
AUTHORS	Chye M.-L. and Xu Z.-F.				
TITLE	Direct Submission				
JOURNAL	Submitted (01-AUG-1999)				
source	Hong Kong, Pokfulam Road, Hong Kong, China				
FEATURES	Location/Qualifiers				
source	1..529				
	/organism="Solanum americanum"				
	/mol_type="mRNA"				
gene	/db_xref="taxon:109975"				
	1..529				
	/gene="PIN2a"				
CDS	5..451				
	/gene="PIN2a"				
	/codon_start=1				
	/product="proteinase inhibitor IIA"				
	/protein_id="AAL36458.1"				
	/db_xref="GI:17221675"				
	/translation="MAVHKVSFLACLLVLGMFLAKHVDKACTRECQFSGYICPR				
	SGSGQKPKICTNCCSGYKCNYSAGKDLICEGSDPRNPDKCTFECDTQIAYSKPCR				
	SGKMKIIRFTGCTTCTCTGCGCYFYDDQDGFVCEGSPKPTATP"				

ORIGIN

Query Match 100.0%; Score 529; DB 8; Length 529;  
Best Local Similarity 100.0%; Pred. No. 1.3e-130;  
Matches 529; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATAATGGCTGTTACAAAGTTAGCTTCCTGTTGCTGCTACTGTTGCTGATGGATGTT 60  
DB 1 CATAATGGCTGTTACAAAGTTAGCTTCCTGTTGCTGCTACTGTTGCTGATGGATGTT 60

QY 61 TCTACTTGGCAACATGTTGATGCCAAGGCTTGCTAGAGATGGTCAATTTAGCTA 120  
DB 61 TCTACTTGGCAACATGTTGATGCCAAGGCTTGCTAGAGATGGTCAATTTAGCTA 120

QY 121 TGGCATATGCCACGTTTCAGAGGAAGTCCCAAAACCTATATGACCAATTTGTTGCTC 180  
DB 121 TGGCATATGCCACGTTTCAGAGGAAGTCCCAAAACCTATATGACCAATTTGTTGCTC 180

QY 181 AGGCTATAAGGGTGGCAACTATTACAGTGTCTAAAGGAGATTTGATTTGTGAAGGAATC 240  
DB 181 AGGCTATAAGGGTGGCAACTATTACAGTGTCTAAAGGAGATTTGATTTGTGAAGGAATC 240

QY 241 TGAACCTAGAAACCCAAAGATTTGTACCTTCGAATGTGATACACAGATTCCTTATTCAAA 300  
DB 241 TGAACCTAGAAACCCAAAGATTTGTACCTTCGAATGTGATACACAGATTCCTTATTCAAA 300

QY 301 ATGTCCTCGTTTCAGAGGAAGATGATAATTAACCCACCTGGATGACCACTTTGTTGCAC 360  
DB 301 ATGTCCTCGTTTCAGAGGAAGATGATAATTAACCCACCTGGATGACCACTTTGTTGCAC 360

QY 361 GGGCTATCAGGGTGTCTACTATTTCGATCAAGATGGTGAATTTGTTCTGTGAAGGAGAG 420  
DB 361 GGGCTATCAGGGTGTCTACTATTTCGATCAAGATGGTGAATTTGTTCTGTGAAGGAGAG 420

QY 421 TCCTGAACCAAGACCACTGCTTATTCTTAATCAATCATATGTTGTTATCTATCAAAAA 480  
DB 421 TCCTGAACCAAGACCACTGCTTATTCTTAATCAATCATATGTTGTTATCTATCAAAAA 480

QY 481 AAATATGATGATGATATATGCTGGTTACTGTAATGAGCTTTATTG 529  
DB 481 AAATATGATGATGATATATGCTGGTTACTGTAATGAGCTTTATTG 529

RESULT 2  
AY422686 687 bp mRNA linear PLN 18-MAY-2004  
LOCUS Solanum nigrum proteinase inhibitor 2b precursor (PIN2b) mRNA,  
DEFINITION complete cds.  
ACCESSION AY422686  
VERSION AY422686.1 GI:40036963  
KEYWORDS Solanum nigrum (black nightshade)  
SOURCE Solanum nigrum  
ORGANISM Solanum nigrum  
REFERENCE 1 (bases 1 to 687)  
AUTHORS Schmidt,D.D., Kessler,A., Kessler,D., Schmidt,S., Lim,M., Gase,K. and Baldwin,I.T.  
TITLE Solanum nigrum: a model ecological expression system and its tools  
JOURNAL Mol. Ecol. 13 (5), 981-995 (2004)  
PUBMED 15078438  
REFERENCE 2 (bases 1 to 687)  
AUTHORS Schmidt,D.D., Gase,K. and Baldwin,I.T.  
TITLE Direct Submission  
JOURNAL Submitted (25-SEP-2003) Molecular Ecology, Max Planck Institute for Chemical Ecology, Hans-Knoell-Strasse 8, Jena D-07745, Germany  
FEATURES Location/Qualifiers  
source  
1..687  
/organism="Solanum nigrum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4112"  
1..687  
/genes="PIN2b"

gene

CDS 28..486  
/gene="PIN2b"  
/codon\_start=1  
/product="proteinase inhibitor 2b precursor"  
/protein\_id="AAR37362.1"  
/db\_xref="GI:40036964"  
/translation="MNVKHEVSSLAYLLVLGLFLFVSAIKHVDKPKTRECNLGVG  
ICPRSGSPENPITCNMGCSGYKNCYSANGTFICGSSDPKNPNTCPYCDGDIAYS  
KCPRSEGQTIIVPTGCTCTCTGYKGYCFKGEFVCEGESIEPNVISNQ"

sig\_peptide 28..120  
/gene="PIN2b"  
/evidence=not\_experimental

ORIGIN  
Query Match 54.4%; Score 288; DB 8; Length 687;  
Best Local Similarity 77.7%; Pred. No. 3.7e-66;  
Matches 363; Conservative 0; Mismatches 95; Indels 9; Gaps 1;

QY 18 AAGTTAGCTTCCTGTTGCTGCTACTGTTGCTTGGATGGATGTTTCTACTTTGCG----- 70  
DB 44 AAGTTAGTTCCCTTGTCTTACTTACTTGTGATTAATTTGTTTATTTGTAAGCGCGA 103

QY 71 --AAACATGTTGATGCAAGGCTTGCTAGAGAAATGTTGCTATTTAGCTATGTCATAT 128  
DB 104 TAAACATGTTGATGCAAGCTTTGACCAAGAAATGTTGCTATTTGGGTATGGAAATAT 163

QY 129 GCCACGTTTCAGAGGAAGTCCCAAAACCTATATGACCAATTTGTTGCTCAGGCTATA 188  
DB 164 GCCCGGTTTCAGAGGAAGTCCGGAANAATCCCAATTTGTTGCTCAGGCTATA 223

QY 189 AGGGTTGCAACTATTACAGTGTCTAAAGGAGATTTGATTTGTGAAGGAGAAATCTGACCCCTA 248  
DB 224 AGGGTTGCAACTATTATTAGTGTCTAAAGGACTTTTATTTGCGAAGGAAGTTCTGACCCCTA 283

QY 249 GAAACCAAAAGATTTGCTTGAATGTGATACAGATGCTTATTTCAAAATGTCCTC 308  
DB 284 AAACCAAAATGCTTCTCTCTATATTGATGAGATATTTGCTATTTCAAAATGTCCTC 343

QY 309 GTTCAGAGGAAGATGATAATTTAAACCCACTGGATGACCACTTTGTCACGGGCTATC 368  
DB 344 GTTCAGAGGAAGAACGTAATATATCCACGGGATGACCACTTTGTCACGGGCTATC 403

QY 369 AGGGTTGCTACTATTTCGATCAAGATGGTGAATTTCTGTGAAGGAGAGATCCTGAAC 428  
DB 404 AGGGTTGCTACTATTATTAGTAAAGAGGTGAGTTTGTGTGAAGGAGAAAGTATTGAAC 463

QY 429 CCAAGACCACTGCTTATTTCTAATCAATCATATGTTGTTATCTATCA 475  
DB 464 CCAACGTTATTCTAATCAATGATATGCGTTGTAGTTTAAATATAA 510

RESULT 3  
LECEVI57G 766 bp mRNA linear PLN 19-JUL-1996  
LOCUS L.esculentum mRNA for protein inhibitor II.  
DEFINITION X94946  
ACCESSION X94946  
VERSION X94946.1 GI:1161571  
KEYWORDS cev157 gene; proteinase inhibitor II.  
SOURCE Lycopersicon esculentum (tomato)  
ORGANISM Lycopersicon esculentum  
REFERENCE 1  
AUTHORS Gadea,J., Mayda,M.E., Conejero,V. and Vera,P.  
TITLE Characterization of defense-related genes ectopically expressed in viroid-infected tomato plants  
JOURNAL Mol. Plant Microbe Interact. 9 (5), 409-415 (1996)  
PUBMED 96252900  
REFERENCE 2 (bases 1 to 766)  
AUTHORS Vera,P.  
TITLE Direct Submission

JOURNAL Submitted (10-JAN-1996) P. Vera, Universidad Politecnica de Valencia, Inst. de Biologia Molecular y Celular de Plantas, Camino de Vera 14, E- 46022 Valencia, SPAIN

FEATURES  
source  
1. .766  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="Rutgers"  
/db\_xref="taxon:4081"  
/issue\_type="leaf"  
1. .766  
/genes="cevi57"  
43. .548  
/genes="cevi57"  
/codon\_start=1  
/product="proteinase inhibitor II"  
/protein\_id="CAA64416.1"  
/db\_xref="GI:1161572"  
/db\_xref="GOA:Q43502"  
/translation="MAVYKVSFLHLLVGMVLLVSTVEHANACTKECNLGVGICPG SEGSPNPICNCCSGKGCNYYANGTFICEGSDPKNPICPSYDQIAYSKCPR SEGKTIYPTGCTCTCYGKCYFQDGEFVCEGSIIPKGTKECDPRVAYMTCP SGLAKLVQVNCSSAGEGCKLYDNGSLTCTGEPQSI STA"  
43. .111  
/genes="cevi57"

gene  
sig\_peptide  
43. .111  
/genes="cevi57"

ORIGIN  
Query Match 54.4%; Score 288; DB 8; Length 766;  
Best Local Similarity 79.2%; Pred. No. 3.6e-66;  
Matches 342; Conservative 0; Mismatches 90; Indels 0; Gaps 0;  
QY 2 ATAATGGCTGTTACAAAGTTAGCTTCCTTGCTGCTACTCTTCTTGGATGGATGTTT 61  
DB 40 ACAATGGCTGTTTACAAAGTTAGTTCTTCTGCTCACCTACTTGTCTTGGAAATGATCTA 99  
QY 62 CTACTTGGCGAATCATGTTGATGCCAAGCTTGTACTAGAGAATGTGGTCAATTTAGCTAT 121  
DB 100 CTAGTAAGCACGGTGGAAACACGCTAATGCTTGTACTAAAGAAATGTGGTAACTCTGGCTAT 159  
QY 122 GGCATATGCCACGTTTCAGAAAGGAAGTCCCAAAACCTATATGCACCAATGTTGCTCA 181  
DB 160 GGGATATGCCACGTTTCAGAAAGGAAGTCCAGAAATTCATATGTACCAATTTGCTCT 219  
QY 182 GGCATATAAGGGTTGCAACTATTACAGTGTCTAAAGGAGATTGATTTGTGAAGGAGATCT 241  
DB 220 GGCATATAAGGGTTGCAACTATTATTACGCTAATGGAACCTTTTATTGTGAAGGACGCTCT 279  
QY 242 GACCCTAGAAACCAAGAAGATTGTACCTTCGATGTGATGATACACAGATTGCTTATCAAAA 301  
DB 280 GATCCAAAAAATCCTAACATTTGCCCTCATATTGTGATCCACAAAATTCCTATTCAAAG 339  
QY 302 TGTCCCTGTTTCAGAAAGAAAGATGATAATTAACCCACCTGGATGCACCACTGTGTGCAGC 361  
DB 340 TGTCCAGTTTCAGAAAGAAAGACGATATCTATCCACAGATGTAGCGTGTGCACT 399  
QY 362 GGCATATCAGGGTTGCTACTATTTTCGATCAAGATGGTGATTTTGTCTGTGAAGGAGAGAT 421  
DB 400 GGTTCAGAGGGTTGCTACTATTTTGGTCAAGATGGAGAGTTTGTGTGTAAGGAGAGAT 459  
QY 422 CCTGAACCCCAAG 433  
DB 460 ATTGAACCTAAG 471

RESULT 4  
STPIN2W  
LOCUS  
DEFINITION S.tuberosum mRNA for pin2 gene, wound induced.  
ACCESSION X99095  
VERSION X99095.1 GI:1431616  
KEYWORDS pin2 gene.  
SOURCE Solanum tuberosum (potato)

ORGANISM Solanum tuberosum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.  
1  
Damann,C., Rojo,E. and Sanchez-Serrano,J.J.  
Abcisic acid and jasmonic acid activate wound-inducible genes in potato through separate organ-specific signal transduction pathways  
Plant J. 11, 101-110 (1997)  
2 (bases 1 to 838)  
Damann,C.  
Direct Submission  
Submitted (04-JUL-1996) C. Damann, Centro Nacional de Biotecnologia, Plant Genetics, Campus de Cantoblanco, 28049 Madrid, Spain  
Location/Qualifiers  
1. .838  
/organism="Solanum tuberosum"  
/macronuclear  
/mol\_type="mRNA"  
/sub\_species="Desiree"  
/db\_xref="taxon:4113"  
/haplotype="tetraploid"  
/cell\_line="Desiree"  
/tissue\_type="root"  
1. .838  
/gene="pin2"  
19. .111  
/gene="pin2"  
/evidence=not\_experimental

Query Match 52.9%; Score 279.6; DB 8; Length 838;  
Best Local Similarity 80.2%; Pred. No. 6.4e-64;  
Matches 356; Conservative 0; Mismatches 79; Indels 9; Gaps 2;  
QY 1 CATATGGCTGTTCACA---AGTTAGCTTCCTTGCTTGCCTACTTGTCTTGGATGGAT 57  
DB 15 CATATGGCTGTTCACAAAGAAAGTTAGTTCCCTTGTCTTACCTACTTGTCTTGGATTAT 74  
QY 58 GTTTCCTACTTGC-----GAAACATCTTGATGTCAAAGGCTTGTACTAGAGAATGTGTCA 111  
DB 75 GGTACTTGTGAAGCGCGATGGACATGTTGATGCGAAGGCTTGTACTTTAGAAATGGTAA 134  
QY 112 TTTTACTATGGCATATGCCACGTTTCAGAAAGAACTCCCAAAAACCTATATATGCACAA 171  
DB 135 TCTTGGCTATGGATATGCCACGTTTCAGAAAGAACTCCGAAAATCCCATATGCACAA 194  
QY 172 TTGTTGCTCAGGCTATAAGGTTTGCACCTATTACAGTGTCTTAAAGGAGATTGATTGTGA 231  
DB 195 CTGTTGTGAGGTTATAAAGGTTGCAATTTATTATAGTCTTAATGGGACTTTTCAATTCGCA 254  
QY 232 AGGAGAACTGACCCCTAGAAACCCAAAGATTGTACTCTTGAATGTGATACACAGATTGC 291  
DB 255 AGGACAACTGACCCCAAAAACCCAAAGCTTGGCCCGGAAATTTGATCCCATATTGC 314  
QY 292 TTATTCAAAATGTCCTCGTTTCAGAAAGAAAGATGATAATTTAAACCCACTGGATGCACAC 351  
DB 315 CTATTCAAAATGTTCCCGTTTCAAGAGGAAAGACACTAAATTTATCCCAACAGATGTACCAC 374  
QY 352 TTGTTGCAGGGCTATCAGGTTTGCCTACTATTTCATCAAGATGGTGTGATTTCCTGTGA 411  
DB 375 ATGTTGCAGGGGTACAAAGGTTGCTACTATTTTGGTAAAGACGGCAAGTTTGTGTTGTGA 434  
QY 412 AGGAGAGAGTCTTGAACCCCAAGAC 435  
DB 435 AGGAGAGATTTTGAACCCCAAGGC 458

RESULT 5  
TOMWIPII  
LOCUS  
DEFINITION Tomato leaf wound-induced proteinase inhibitor II mRNA, complete  
TOMWIPII 660 bp mRNA linear PLN 27-APR-1993  
DEFINITION



```

RESULT 7
STRIN2G
LOCUS STPRIN2G 1695 bp DNA linear PLN 11-MAY-1995
DEFINITION S.tuberosum gene for proteinase inhibitor II.
ACCESSION 212753
VERSION 212753.1 GI:21553
KEYWORDS proteinase inhibitor; proteinase inhibitor II.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 1695)
Choi,Y., Moon,Y. and Lee,J.S.
Primary structure of two proteinase inhibitor II genes closely
linked in the Potato genome
Hanguk Saenghwahakhoe Chi 23, 214-220 (1990)
2 (bases 1 to 1695)
Lee,J.S.
Direct Submission
Submitted (19-JUN-1992) Jong S. Lee, Molecular Biology, Seoul
National University, ShinRim Dong, Kwanak-Gu, Seoul, 151-742,
Republic of Korea
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. .1695
/organism="Solanum tuberosum"
/mol_type="genomic DNA"
/cultivar="Russel Burbank"
/db_xref="taxon:4113"
repeat_region
40. .148
/notes="partial long direct repeat unit 1"
repeat_region
262. .378
/notes="partial long direct repeat unit 2"
CAAT_signal
438. .441
TATA_signal
545. .550
CDS
join(629. .677,779. .1206)
/codon_start=1
/product="proteinase inhibitor II"
/protein_id="CAA78277.1"
/db_xref="GI:21554"
/db_xref="GOA:Q00782"
/db_xref="Swiss-Prot:Q00782"
/translation="MAIHKVSVFLAYLLVGLMLFVSMHVDAKACTLECGNLGYGI
CPRSEGPENICNCAGYKGNYSANGTFICEGQSHPKPKACPRNCDPHIAYSK
CPRSGGKTLIYPTGCTTCCTGYTDCYVFGKDGKFCVCEGSEIEPKACTLECDRLHT"
629. 677
/number=1
exon
678. .778
intron
779. .1206
exon
/number=2
polyA_signal
1476. .1481
misc_structure
1562. .1568
/notes="mRNA 3' maturation signal"
ORIGIN
Query Match 49.8%; Score 263.6; DB 8; Length 1695;
Best Local Similarity 82.5%; Pred. No. 1.2e-59;
Matches 302; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
QY 70 GAAACATGTTGATGCCAAGCTTGTTACTAGAGAATGGTCAATTTTAGCTATGGCATATG 129
DB 804 GGAACATGTTGATGCCAAGCTTGTTACTTTAGAATGGTAAATCTTGGCTATGGCATATG 863
QY 130 CCACGTTGACAGGAAGTCCCAAAACCTATATGCACCAATTTGCTCAGGCTATAA 189
DB 864 CCCACGTTGACAGGAAGTCCGAAAATCCCATATGCACCACTGTTGTGACGTTATAA 923
QY 190 GGGTTGCAACTATTACAGTCTCTAAAGGAGATTTGATTTGTGAAGGAGAACTCTGACCTAG 249
DB 924 AGTTGCAACTATTATAGTCTTAATGGACTTTTATTTTGTGAAGGACAATCTCACCANA 983

```

```

QY 250 AAACCCAAAAGATTGTACCTTCGAATGTGATACACAGATTGCTTATTCAAATGTCTCG 309
DB 984 AAACCCAAAAGCTTGCCCGAATTTGTATCCACATATTGCCCTATTCAAATGTCCACG 1043
QY 310 TTCAGAAAGAAAGATGATAATTAACCCACTGTGATGCCACCTTGTTCACGGGCTATCA 369
DB 1044 TTCAGAGGAAAGACCTTAATTTATCCACAGGATGTACCACATGTTGCACGGGTACAC 1103
QY 370 GGGTTCCTACTATTTCGATCAAGATGGTGAATTTTGTCTGTGAAGGAGAGAGTCTTGAACC 429
DB 1104 GGATTCTCTATTATTGTTGTAAGACGGCAAGTTTGTGTGTGAAGGAGAGATATTGAACC 1163
QY 430 CAAGAC 435
DB 1164 CAAGGC 1169

STU45450 2068 bp DNA linear PLN 27-JAN-1999
LOCUS Solanum tuberosum proteinase inhibitor II (pin2r) gene, complete cds.
DEFINITION
ACCESSION U45450
VERSION U45450.1 GI:1206014
KEYWORDS
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 2068)
Park,S. and Thornburg,R.
Isolation and characterization of a proteinase inhibitor II gene
that is not wound-inducible (Accession No. U45450) (PGR96-007)
Plant Physiol. 110, 1048 (1996)
2 (bases 1 to 2068)
Thornburg,R.W. and Park,S.
Direct Submission
Submitted (11-JAN-1996) Robert W. Thornburg, Biochemistry and
Biophysics, Iowa State University, Ames, IA 50011, USA
JOURNAL
FEATURES
Location/Qualifiers
source
1. .2068
/organism="Solanum tuberosum"
/mol_type="genomic DNA"
/strain="Russel Burbank"
/db_xref="taxon:4113"
1014. .1917
/gene="pin2"
join(1014. .1109,1223. .1917)
/gene="pin2"
join(1058. .1109,1223. .1614)
/gene="pin2"
/function="inhibitor of trypsin and chymotrypsin"
/codon_start=1
/product="proteinase inhibitor II"
/protein_id="AAD09849.1"
/db_xref="GI:1206015"
/translation="MAVHKVSVFVAYLLVGLMFLYVDALGCTKECNLGFICPRSE
GSPTNPICNCCSGYKGNYSAGFLICQGESDPKNPKACPLNCDTIAYSCRPRSE
GKSLIYPTGCTTCCTGCTGKCYGFGINGKFCVCEGSEDEPKPYMTA"
ORIGIN
Query Match 49.3%; Score 260.6; DB 8; Length 2068;
Best Local Similarity 82.4%; Pred. No. 7.4e-59;
Matches 299; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
QY 73 ACATGTTGATGCCAAGCTTGTTACTAGAGAATGGTCAATTTTAGCTATGGCATATGCC 132
DB 1233 ATATGTTGATGCTTTGGTGTGTTACTAAAGAATGTGTAATCTTGGCTTTGGGATATGCC 1292
QY 133 AGTTTCAGAGGAAGTCCCAAAACCTATATGCACCAATTTGCTCAGGCTATAAGGG 192
DB 1293 AGTTTCAGAGGAAGTCCCAAAACCTATATGCACCAATTTGCTCAGGCTATAAGGG 1352

```



```
Matches 327; Conservative 0; Mismatches 113; Indels 0; Gaps 0;
QY 5 ATGGCTGTTCCACAAAGTTAGCTTCCTTGGCTGCTACTGTTCTTGGATGGATGTTTCTA 64
Db 112 ATTGGTTTATACATACACAAAGTAGTTTATATATTTTCTTATATATATTTGTTGAGGA 171
QY 65 CTTCGCAACATGCTTATGATCCCAAGGCTTGCTAGAGAAATGCTGCTATTTAGCTATGCC 124
Db 172 ATGTTTCTATATGTTGATCCCAAGGCTTGCTAGAGAAATGCTGCTATTTAGCTATGCCG 231
QY 125 ATATGCCCAAGCTTCCCAAGAAAGTCCCAAAACCTATATGCAACCAATTTGTTCTCAGGC 184
Db 232 ATATGCCCAAGCTTCCCAAGAAAGTCCGCTAAATCCCATATGCAATTTGTTCTCAGGC 291
QY 185 TATAAGGTTGCAACTATTACAGTGTAAAGAGATTTGATTTGTGAAGAGAAATCTGAC 244
Db 292 TATAAGGTTGTAATATTATTAATTTCTTCGGAATAATTTATTTGTGAAGAGAAATCTGAT 351
QY 245 CCTAGAAACCCCAAGATTTGCTGCAATGATGATACAGATTTGCTTATTCAGAAATGT 304
Db 352 CCAAAAGGCAATGCTTGCACCTTTAATTTGTGATCCAAATATTTGCTTATTCAGAAATGT 411
QY 305 CTCTGTTTCAGAAAGAAAGATGATAATTAACCCACTGGATGCAACCACTTTGTTGCAGGGC 364
Db 412 CCCCCTTTCAGAAAGAAATGCTTAAATTTATCCACCGGATGATACCAAGCTTTGCACAGG 471
QY 365 TATCAGGTTGCTATTTTCGATCAAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 424
Db 472 TACAAGGTTGCTATTTATTTTGGTAAAGATGGAAGTTTGTATGTTGAAGAGAGAGTGTAT 531
QY 425 GAACCAAGCACTGCTTA 444
Db 532 GAACCAAGCAATATGTA 551

RESULT 11.
AY129402
LOCUS AY129402 1670 bp DNA linear PLN 16-JUN-2004
DEFINITION Lycopersicon esculentum proteinase inhibitor II gene, complete cds.
ACCESSION AY129402
VERSION AY129402.1 GI:33413549
KEYWORDS Lycopersicon esculentum (tomato)
SOURCE Lycopersicon esculentum
ORGANISM Lycopersicon esculentum
REFERENCE 1 (bases 1 to 1670)
AUTHORS Zhang,H.Y., Xie,X.Z., Xu,Y.Z. and Wu,N.H.
TITLE Isolation and functional assessment of a tomato proteinase inhibitor II gene
JOURNAL Plant Physiol. Biochem. 42 (5), 437-444 (2004)
PUBMED 15191748
REFERENCE 2 (bases 1 to 1670)
AUTHORS Zhang,H., Xie,X., Wu,N. and Huang,M.
TITLE Direct Submission
JOURNAL Submitted (08-JUL-2002) Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Nanyitiao No.3 Zhongguancun, Beijing 2707, China
FEATURES
source
1. .1670
/organism="Lycopersicon esculentum"
/mol_type="genomic DNA"
/db_xref="taxon:4081"
join(-995..1046..1156..>1550)
/product="proteinase inhibitor II"
join(995..1046,1156..1550)
/codon_start=1
/product="proteinase inhibitor II"
/protein_id="AA05016.1"
/db_xref="GI:33413550"
translation="MAVHKVNFVAYLLVLGMFLYVDKACTRECNGLFGICPRSE
GSPLPINCCSGYKGCYNNYSFGKFCIEGESDPRKPNACTFNCDPNIAYSRCPRSE
```

```
ORIGIN
Query Match 49.0%; Score 259.2; DB 8; Length 1670;
Best Local Similarity 74.3%; Pred. No. 1.8e-58;
Matches 327; Conservative 0; Mismatches 113; Indels 0; Gaps 0;
QY 5 ATGGCTGTTCCACAAAGTTAGCTTCCTTGGCTGCTACTGTTCTTGGATGGATGTTTCTA 64
Db 1098 ATTGGTTTATACATACACAAAGTAGTTTATATATTTTCTTATATATATTTGTTGAGGA 1157
QY 65 CTTCGCAACATGTTGATCCCAAGGCTTGCTAGAGAAATGCTGCTATTTAGCTATGCC 124
Db 1158 ATGTTTCTATATGTTGATCCCAAGGCTTGCTAGAGAAATGCTGCTATTTAGCTATGCCG 1217
QY 125 ATATGCCCAAGCTTCCCAAGAAAGTCCCAAAACCTATATGCAACCAATTTGTTCTCAGGC 184
Db 1218 ATATGCCCAAGCTTCCCAAGAAAGTCCGCTAAATCCCATATGCAATCAATTTGTTCTCAGGC 1277
QY 185 TATAAGGTTGCAACTATTACAGTGTAAAGAGATTTGATTTGTGAAGAGAAATCTGAC 244
Db 1278 TATAAGGTTGTAATATTATTAATTTCTTCGGAATAATTTATTTGTGAAGAGAAATCTGAT 1337
QY 245 CCTAGAAACCCCAAGATTTGCTGCAATGATGATACAGATTTGCTTATTCAGAAATGT 304
Db 1338 CCAAAAGGCAATGCTTGCACCTTTAATTTGTGATCCAAATATTTGCTTATTCAGATGT 1397
QY 305 CTCTGTTTCAGAAAGAAAGATGATAATTAACCCACTGGATGCAACCACTTTGTTGCAGGGC 364
Db 1398 CCCCCTTTCAGAAAGAAATCGTTAAATTTATCCACCGGATGATACCAAGCTTTGCACAGG 1457
QY 365 TATCAGGTTGCTATTTTCGATCAAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 424
Db 1458 TACAAGGTTGCTATTTATTTTGGTAAAGATGGAAGTTTGTATGTTGAAGAGAGAGTGTAT 1517
QY 425 GAACCAAGCACTGCTTA 444
Db 1518 GAACCAAGCAATATGTA 1537

RESULT 12
POTIKA
LOCUS POTIKA 1241 bp DNA linear PLN 28-APR-1994
DEFINITION S.tuberosum (potato) inhibitor II-chloramphenicol acetyltransferase gene (IIK), complete cds.
ACCESSION M15186
VERSION M15186.1 GI:169488
KEYWORDS inhibitor II-chloramphenicol acetyl transferase.
SOURCE Solanum tuberosum
ORGANISM Solanum tuberosum
REFERENCE 1 (bases 1 to 1241)
AUTHORS Thornburg,R.W., An,G., Cleveland,T.E., Johnson,R. and Ryan,C.A.
TITLE Wound-inducible expression of a potato inhibitor II-chloramphenicol acetyltransferase gene fusion in transgenic tobacco plants
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 84, 744-748 (1987)
COMMENT Original source text: Solanum tuberosum DNA.
FEATURES
source
1. .1241
/organism="Solanum tuberosum"
/mol_type="genomic DNA"
/db_xref="taxon:4113"
prim_transcript 258..1042
258..357
/note="inhibitor IIK mRNA and introns"
exon
/gene="IIK"
number=1
join(306..357,475..884)
CDS
/gene="IIK"
/note="precursor"
/codon_start=1
```









**THIS PAGE IS BLANK**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2005, 04:42:11 ; Search time 2581 Seconds  
(without alignments)  
7468.663 Million cell updates/sec

Title: US-10-725-829-1  
Perfect score: 529  
Sequence: 1 cataatggctgttcacaag.....ctgtaatggactttattg 529

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Databae : EST:\*

1: gb\_est1: \*  
2: gb\_est2: \*  
3: gb\_hic: \*  
4: gb\_est3: \*  
5: gb\_est4: \*  
6: gb\_est5: \*  
7: gb\_est6: \*  
8: gb\_gss1: \*  
9: gb\_gss2: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	362.8	68.6	545	4	BI176253
2	362.8	68.6	554	2	BF053417
3	362.8	68.6	571	5	Q114991
4	337.4	63.8	451	5	Q114990
5	289.6	54.7	469	2	AW035333
6	289.6	54.7	614	4	BI421156
7	289.6	54.7	626	4	BI205417
8	289.6	54.7	663	4	BI205603
9	270.8	51.2	625	4	BI433294
10	270.8	51.2	628	4	BI433271
11	269.4	50.9	672	2	BF133853
12	269.4	50.9	675	4	BI436259
13	269.4	50.9	677	2	BF153879
14	267.8	50.6	591	2	BF153508
15	267.2	50.5	779	4	BI421162
16	267	50.5	641	2	BF460374
17	267	50.5	655	4	BI406670
18	265.4	50.2	477	2	AW093009
19	265.4	50.2	484	2	BE463283
20	265.4	50.2	501	1	AI771251
21	265.4	50.2	503	2	AW221025
22	265.4	50.2	504	1	AI485972
23	265.4	50.2	515	1	AI489819
24	265.4	50.2	533	1	AI771306

25	265.4	50.2	550	1	AI485979
26	265.4	50.2	558	2	AW092640
27	265.4	50.2	564	2	BE354745
28	265.4	50.2	571	1	AI771944
29	265.4	50.2	581	4	BI929125
30	265.4	50.2	582	2	BE354912
31	265.4	50.2	587	1	AI484805
32	265.4	50.2	604	1	AI487479
33	265.4	50.2	612	2	BF460196
34	265.4	50.2	623	4	BI930668
35	265.4	50.2	623	4	BI930919
36	265.4	50.2	623	4	BI932864
37	265.4	50.2	624	2	AW093425
38	265.4	50.2	626	4	BI929471
39	265.4	50.2	629	2	AW217790
40	265.4	50.2	630	4	BI931013
41	265.4	50.2	637	2	BE463161
42	265.4	50.2	644	4	BI935830
43	265.4	50.2	645	4	BG125289
44	265.4	50.2	655	4	BI930423
45	265.4	50.2	655	4	BI933036

## ALIGNMENTS

RESULT 1  
BI176253  
LOCUS EST517336 cSTS Solanum tuberosum cDNA clone cSTS28K7 5' sequence, linear EST 07-MAR-2003  
DEFINITION mRNA sequence.  
ACCESSION BI176253  
VERSION BI176253.1 GI:14642064  
KEYWORDS EST.  
SOURCE Solanum tuberosum (potato)  
ORGANISM Solanum tuberosum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.  
REFERENCE 1 (bases 1 to 545)  
AUTHORS van der Hoeven, R., Bezzerides, J., Sun, H., Cho, J., Chiemingo, A., Bougri, O., Buell, C.R., Ronning, C., Tankley, S. and Baker, B.  
TITLE Generations of ESTs from sprouting potato eyes  
JOURNAL Unpublished (2000)  
COMMENT Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL:  
http://genome.arizona.edu/orders/  
Seq primer: M13P-R.

FEATURES  
source location/Qualifiers  
1..545  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="cSTS28K7"  
/tissue\_type="sprouting eyes from tubers"  
/dev\_stage="12-14 weeks post harvest"  
/lab\_host="SOLR"  
/clone\_lib="cSTS"  
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; Various sizes of sprouting eyes (2mm to 15mm) were taken from tubers. The tubers were incubated at 26C in the dark for 2-3 weeks prior to sprouting. The eyes were frozen in liquid nitrogen immediately upon removal from tubers."

## ORIGIN

Query Match 68.6%; Score 362.8; DB 4; Length 545;  
Best Local Similarity 84.6%; Pred. No. 4.4e-89;



9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
This clone can be obtained from the University of Arizona Genomics  
Institute. Orders can be made through URL:  
http://genome.arizona.edu/orders/  
Seq primer: T7

**FEATURES**  
**source**  
Location/Qualifiers  
1..571  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec or Binjete"  
/db\_xref="taxon:4113"  
/clone="STMX74"  
/tissue\_type="mixed tissues"  
/lab\_host="SOLR"  
/clone\_lib="mixed potato tissues"  
/note="vector: pBluescript SK(-); Site 1: EcorI; Site 2:  
XhoI; supplier: Combination of untreated and Phytophthora  
infestans-created libraries of stolons, leaves, leaflets  
axillary buds of stem explants, petioles, germinating  
eyes, tubers, or roots."

**ORIGIN**

Query Match	68.6%;	Score 362.8;	DB 5;	Length 571;
Best Local Similarity	84.6%;	Pred. No. 4.5e-89;		
Matches 445;	Conservative 0;	Mismatches 72;	Indels 9;	Gaps 3;

```

QY      1  CATAATGGCTGTTCCACAAAGTTAGCTTCCTTGCTTGGCTACTTGTCTCTTGGATGGATGTT  60
DB      552  CATATGCGCTGCTCACAAGTTAGCTTCATTCCTTGGCTACTTGTCTCTTG---GGATGTT  496

QY      61  TCTACTTGCAGAAACATGTTGATGCCAAGGCTTGTA CTAGAGAAATGTGTGTCATTTTAGCTA  120
DB      495  TCTTCTTTGTAACACACTGTTGATGCCAAGGTGTGTC CAAGATATGTGTGTAATTTTGGGTA  436

QY      121  TGGCATATGCCAGCTTCAGAGGAAGTCCCCAAAACCTATATGCACCAATTTGTTGCTC  180
DB      435  TGGACTATGCCACGTTCTGAAGGAAGTCCGGNAAAGCCCATATGCACCAATTTGTTGCTC  376

QY      181  AGGCTATAAGGGTTGCAACTATTACAGTGTCTAAAGGAGATTGATTTGTGGAAGGAGAATC  240
DB      375  GGGCTACAGGGTTGCAATATTAACAGTGCTAACGGAGATTAGTTGTGGAAGGGAGCTC  316

QY      241  TGACCCCTAGAAACCCCAAAGATTGTACCTCCGAATGTGATACAGAGTTGCTTATTCAAA  300
DB      315  TGACCCCGAGAAACCCAAATGATTGTTCTATGAGTGTGATACACAAATTGCTATTCAA  256

QY      301  ATGTCTCTGTTCAGAAGGAAAGATGATAATTAAACCCACTGGATGCACCACTTGTTCAC  360
DB      255  ATGTCTCTGTTCTGAAGGAAGGATGATAAATTAACCCCAACGGGTTGTACCAATGTTCAC  196

QY      361  GGGCTATCAGGTTGCTACTATTTCGATCAAGATGGTGATTTTCTGTGTAAGGAGAGAG  420
DB      195  CGGCTACAGGTTGCTACTATTTCGATCAAGAAGGTGATTTTGTGTGTAAGGAGAGAG  136

QY      421  TCCTGGAACCAAGACCACTGCT----TATTTCTAATCAATCATATGTTGTTATCTATCAA  476
DB      135  TCCTGGAACCAAGCAGTGGCTTAAGTATTTCCTAATTAATCGTTGTGCAATTATTAG  76

QY      477  AAAAAAATATGATGCATGATATATGCTGGTTACTGTGAATGTGCAC  522
DB      75  TAACAA--ATGTGTGCATTATATATGCTAATCTTTTGTAAACGTGCAC  32

```

**RESULT 4**  
BQ114990  
LOCUS  
DEFINITION  
EST600566 mixed potato tissues Solanum tuberosum cDNA clone STMX74  
5', end, mRNA sequence.  
ACCESSION  
BQ114990  
VERSION  
BQ114990.2 GI:21916736  
KEYWORDS  
EST.  
Solanum tuberosum (potato)  
SOURCE

**ORGANISM**

Solanum tuberosum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamiales; Solanaceae; Solanum.

**REFERENCE**

Buell, C.R., Hart, A., Baker, B., Tanksley, S., Fry, W., Smart, C.,  
Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J. and  
Karanycheva, S.A.  
Generation of a set of potato cDNA clones for microarray analyses  
Unpublished (2002)  
On Apr 17, 2002 this sequence version replaced gi:2016952.

**TITLE**

Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@igr.org  
This clone can be obtained from the University of Arizona Genomics  
Institute. Orders can be made through URL:  
<http://genome.arizona.edu/orders/>

**FEATURES**

Seq primer: T3.	Location/Qualifiers
	1..451
	/organism="Solanum tuberosum"
	/mol_type="mRNA"
	/cultivar="Kennebec or Binjte"
	/db_xref="taxon:4113"
	/clone="STMCX74"
	/tissue type="mixed tissues"
	/lab host="SOLR"
	/clone_lib="mixed potato tissues"
	/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
	XhoI; supplier: Combination of untreated and Phytophthora
	infestans-treated libraries of stolons, leaves, leaflets,
	axillary buds of stem explants, petioles, germinating
	eyes, tubers, or roots."

**ORIGIN**

Query Match	63.8%; Score 337.4; DB 5; Length 451;
Best Local Similarity	87.6%; Pred.No. 4.3e-82;
Matches 381;	Conservative 0; Mismatches 51; Indels 3; Gaps 1;

```

Qy      1   CATAAATGGCTGTTCAAAAAGTTAGCTTCCTTGCTTGCCTACTGTCTTGGATGGATGTT 60
Db      20   CATAAATGGCTGCTCACAAGAATTAGCTTCATTTGCTTGCCTACTGTCTTCTTCTT 76

Qy     61   TCTACTTGGCAACATGTTTGATGCCAAGGCTTCTACTAGAGAATGGTTCATTTTAGCTA 120
Db     77   TCTTCTTGTAACACATGTTTGATGCCAAGGTGTGTCCAAGAATATGTGGTAATTTTGGGA 136

Qy    121   TGCCATATGCCACGTTTCAGAGGAAGTCCCCAAAAACCCTATATGCACCAATTTGTGCTC 180
Db    137   TGGACTATGCCACGTTCTGAAGGAAGTCGGGAAGCCCATATGCACCAATTTGTGCTC 196

Qy    181   AGGCTATAAGGGTTGCAACTATTACAGTGCTAAAGAGAGATTGTGATTGTGAAGAGAAATC 240
Db    197   GGCGTCAAGGGTTGCAAAATATTACAGTGCTAACGGAGATTAGTTGTTGTGAAGGGAGTC 256

Qy    241   TGACCCTAGNAACCCCAAAGATTGTACCTTCGAATGTGATACACAGATTGCTATTTCAAA 300
Db    257   TGACCCCAAGAAACCCCAATGATTGTTCCTATGATGTGATACACAAATTCGCCATTTCAAA 316

Qy    301   ATGTCCTCGTTTCAGAGGAAAGATGATAATTAACCCACCTGGATGCACCACTTTGTGTCAC 360
Db    317   ATGTCCTCGTTCTGAAGGAAGGATGATATTAACCCACCGGTTGTACCAATTTGTGTCAC 376

Qy    361   GGGCTATCATGGGTTGCTACTATTTCGATCAAGATGGTGAATTTTTGTCTGTGAAGGAGAG 420
Db    377   CGGCTACCGGGTTGCTACTATTTTCGATCAAGAAGGTGATTTTGTGTGTGAAGGAGAG 436

Qy    421   TCCTGAACCCCAAGAC 435
Db    437   TCCTGAACCCCAAGCC 451
  
```

[illegible]

```

RESULT 5
AW035333
LOCUS
DEFINITION
  EST280696 tomato callus, TAMU Lycopersicon esculentum cDNA clone
  cLEC40H22 similar to protein inhibitor II, mRNA sequence.
ACCESSION
AW035333
VERSION
AW035333.1 GI:5894089
SOURCE
  Lycopersicon esculentum (tomato)
ORGANISM
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE
  1 (bases 1 to 469)
  Alcalá,J., Vrebalov,J., White,R., Matern,A.L., Vision,T.,
  Holt,I.E., Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ahn,S.,
  Ronning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and
  Giovannoni,J.
  Generation of ESTs from tomato callus tissue
  Unpublished (1999)
  Contact: CUGI
  Clemson University Genomics Institute
  Clemson University
  100 Jordan Hall, Clemson, SC 29634, USA
  Email: http://www.genome.clemson.edu/orders/index.html
  5 prime sequence.
FEATURES
  source
    Location/Qualifiers
      1..469
        /organism="Lycopersicon esculentum"
        /mol_type="mRNA"
        /cultivar="TA496"
        /db_xref="taxon:4081"
        /clone="cLEC40H22"
        /tissue_type="callus"
        /dev_stage="25-40 days old"
        /lab_host="XLI-Blue MRF"
        /clone_lib="tomato callus, TAMU"
        /note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
        XhoI; supplier: Giovannoni laboratory; cLEC - Cotyledons
        of seedlings 7-10 days post-germination were excised, cut
        at both ends and placed on MS medium with no selection.
        Mixed callus was harvested at 25 and 40 days and included
        undifferentiated masses. Tomato Callus EST Library"
ORIGIN
  Query Match          54.7%; Score 289.6; DB 2; Length 469;
  Best Local Similarity 79.4%; Pred. No. 6.9e-69;
  Matches 343; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY  2  ATAAATGGCTGTTTACAAAGTTAGCTTCCTTGGCTTGCTACTTCTTGGATGGATGTTT 61
DB  32 ACATAGGCTGTTTACAAAGTTAGCTTCCTTGGCTTGCTACTTCTTGGATGATCTA 91

QY  62 CTACTTGCAGAACATGTTGATGCCAAGGCTTGTAAGAGAAATGTGGTCAATTTAGCTAT 121
DB  92 CTAGTAAGCACGCTGGAACACGCTAATGCTTGTAACAAAGATGTGGTAATCTTGGCTAT 151

QY  122 GGCATATGCCACGTTTCAGAGGAAGTCCCAAAAACCTATATGCAACCAATGTTGCTCA 181
DB  152 GGGATATGCCACGTTTCAGAGGAAGTCCAGAAAATCCAATATGTACCAATTTGCTCT 211

QY  182 GGCATATAAGGGTTGCAACTATTATACAGTGTCTTAAAGGAGATTGTTGTTGAAGGAGATCT 241
DB  212 GGCCTAATAGGGTTGCAACTATTATTAACGCTAATGGAACITTTTATTTGTGAAGGAGCTCT 271

QY  242 GACCCCTAGAAACCCAAAGATTGTACCTTCGAATGTGATACACAGATTTGCTATTCAAAA 301
DB  272 GATCCAAAAAATCCTAACATTTGCCCTCATATTGTGATCCACAAATTTGCTATTCAAAG 331

QY  302 TGTCTTCGTTTCAGAGGAAGATGATAATTAACCCCACTGGATGCAACCACTTGTTCACG 361
DB  332 TGTCCACGTTTCAGAGGAAGACGATAATCTATATCCCAAGGATGTACGACGTTGTGCACT 391

```

---

```

QY  362 GGCTATCAGGGTTGCTACTATTTCGATCAAGATGGTGATTTTCTCTGTGAAGGAGAGT 421
DB  392 GGGTACAAAGGGTTGCTACTATTTCGATCAAGATGGAGAGTTTCTGTGTAAGGAGAGT 451

QY  422 CTGGAACCCCAAG 433
DB  452 ATTGAACCTTAAG 463

RESULT 6
BI421156
LOCUS
DEFINITION
  EST531822 tomato callus, TAMU Lycopersicon esculentum cDNA clone
  cLEC66G19 5' end, mRNA sequence.
ACCESSION
BI421156
VERSION
BI421156.1 GI:15194430
KEYWORDS
SOURCE
  Lycopersicon esculentum (tomato)
ORGANISM
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE
  1 (bases 1 to 614)
  Alcalá,J., Vrebalov,J., White,R., Matern,A.L., Vision,T.,
  Holt,I.E., Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ahn,S.,
  Ronning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and
  Giovannoni,J.
  Generation of ESTs from tomato callus tissue
  Unpublished (1999)
  Contact: CUGI
  Clemson University Genomics Institute
  Clemson University
  100 Jordan Hall, Clemson, SC 29634, USA
  Email: http://www.genome.clemson.edu/orders/index.html.
FEATURES
  source
    Location/Qualifiers
      1..614
        /organism="Lycopersicon esculentum"
        /mol_type="mRNA"
        /cultivar="TA496"
        /db_xref="taxon:4081"
        /clone="cLEC66G19"
        /tissue_type="callus"
        /dev_stage="25-40 days old"
        /lab_host="XLI-Blue MRF"
        /clone_lib="tomato callus, TAMU"
        /note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
        XhoI; supplier: Giovannoni laboratory; cLEC - Cotyledons
        of seedlings 7-10 days post-germination were excised, cut
        at both ends and placed on MS medium with no selection.
        Mixed callus was harvested at 25 and 40 days and included
        undifferentiated masses. Tomato Callus EST Library"
ORIGIN
  Query Match          54.7%; Score 289.6; DB 4; Length 614;
  Best Local Similarity 79.4%; Pred. No. 7.3e-69;
  Matches 343; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY  2  ATAAATGGCTGTTTACAAAGTTAGCTTCCTTGGCTTGCTACTTCTTGGATGGATGTTT 61
DB  13 ACAATGGCTGTTTACAAAGTTAGTTTCTTGGCTTGCTACTTCTTGGATGATCTA 72

QY  62 CTACTTGCAGAACATGTTGATGCCAAGGCTTGTAAGAGAAATGTGGTCAATTTAGCTAT 121
DB  73 CTAGTAAGCACGCTGGAACACGCTAATGCTTGTAACAAAGATGTGGTAATCTTGGCTAT 132

QY  122 GGCATATGCCACGTTTCAGAGGAAGTCCCAAAAACCTATATGCAACCAATGTTGCTCA 181
DB  133 GGGATATGCCACGTTTCAGAGGAAGTCCAGAAAATCCAATATGTACCAATTTGCTCT 192

QY  182 GGCATATAAGGGTTGCAACTATTATACAGTGTCTTAAAGGAGATTGTTGTTGAAGGAGATCT 241
DB  193 GGCTAATAGGGTTGCAACTATTATTAACGCTAATGGAACITTTTATTTGTGAAGGAGAGCTCT 252

```



Db 28 ACATGGCTTTACAAAGTAGTTCCCTTGCTCACCCTACTTCTTCTTGGAAATGTATCTA 87  
 QY 62 CTACTTCCGCAACATGTTGATGCCAAGGCTGTGACTAGAGAAATGTGGTCAATTTAGCTAT 121  
 Db 88 CTAGTAAGCAGCGTGGAAACAGCGCTAATGCTTGTACTAAAGAAATGGTAAATCTTGGCTAT 147  
 QY 122 GGCATATGCCACGTTTCAGAAAGAGTCCCAAAAACCTATATGACCAAAATGTTGGCTCA 181  
 Db 148 GGGATATGCCAGGTTTCAGAAAGAGTCCAGAAAAATCCAAATATGTATGTAATGTTGCTCT 207  
 QY 182 GGCATATAGGGTTGCCAATATTACAGTCTTAAGGAGATTTGATTGTTGAAGGAGATCT 241  
 Db 208 GGCATATAGGGTTGCCAATATTACAGTCTTAAGGAGATTTGATTGTTGAAGGAGATCT 267  
 QY 242 GACCTAGAAACCCAAAGATTTGACCTTCGAATGTGATACACAGATTTGCTTATTCAAA 301  
 Db 268 GATCCAAAATCTTACATTTGCCCTCATATTTGATGCCAAATGCTTATTCAGAG 327  
 QY 302 TGTCTCTGTTTCAGAAAGAGATGATAATTAACCCACCTGGATGACCACTTGTGGACG 361  
 Db 328 TGTCCACGTTTCAGAAAGAGAGATAATCTATCCACAGGATGTACGAGCTGTGCACT 387  
 QY 362 GGCATATAGGGTTGCTACTATTTTCGATCAGATGGTGTATTTCTCTGTAAGGAGAGT 421  
 Db 388 GGTTCAGAAAGGTTGCTACTATTTTGGTCAAGATGGAGAGTTTGTGTGAAGGAGAGT 447  
 QY 422 CCTGAACCCCAAG 433  
 Db 448 ATTGAACCTTAAG 459

RESULT 9  
 BI433294  
 LOCUS  
 DEFINITION  
 EST536055 P. infestans-challenged potato leaf, compatible reaction  
 Solanum tuberosum cDNA clone PPCBC78 5' sequence, mRNA sequence.  
 BI433294  
 BI433294.1 GI:15257984  
 EST.  
 Solanum tuberosum (potato)  
 Solanum tuberosum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamiales; Solanales; Solanaceae; Solanum.  
 1 (bases 1 to 625)  
 Restrepo,S., Griffiths,H.M., Smart,C.D., Cho,J., Chiemiango,A.,  
 Bougri,O., Buell,C.R., Ronning,C.M., Fry,W.E. and Baker,B.  
 Generation of ESTs from Potato Leaves Challenged with Phytophthora  
 infestans, Compatible Interaction  
 Unpublished (2000)  
 Contact: Robin Buell  
 The Institute for Genomic Research  
 9712 Medical Center Dr, Rockville, MD 20850, USA  
 Email: potato-array@tigr.org  
 This clone can be obtained from the University of Arizona Genomics  
 Institute. Orders can be made through URL:  
 http://genome.arizona.edu/orders/  
 Seq primer: M13P-R.

FEATURES  
 source  
 1..625  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /cultivar="Kennebec"  
 /db\_xref="taxon:4113"  
 /clone="PPBC78"  
 /tissue\_type="leaf"  
 /dev\_stage="6 week old"  
 /lab\_host="SOLR"  
 /clone\_lib="P. infestans-challenged potato leaf,  
 compatible reaction"  
 /notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; supplier: Cornell University, Fry lab; sequencing:  
 The Institute for Genomic Research; Whole plants were

challenged with 20,000 sporangia/ml of the compatible P.  
 infestans isolate US 940480. Leaf tissue was collected at  
 3, 6, 9, 12, 24, 48, 72 hours after inoculation and frozen  
 in liquid nitrogen immediately upon removal. Kennebec  
 plants showed first symptoms of infection at 48 hours  
 after inoculation. NOTE: We cannot exclude the  
 possibility that this sequence is actually derived from  
 Phytophthora rather than potato."

## ORIGIN

Query Match 51.2%; Score 270.8; DB 4; Length 625;  
 Best Local Similarity 79.7%; Pred. No. 1.1e-53;  
 Matches 349; Conservative 0; Mismatches 77; Indels 12; Gaps 2;  
 QY 1 CATATGGCTGTTCACA---AAGTTAGCTTCCCTTGCCTACTTGTCTTGGATGGAT 57  
 Db 23 CATATGGCTGTTCACAAGGAAGTTAGTTTCGTTGCTTACCTACTAATTTGTTCTTGAAT 82  
 QY 58 GTTTCCTACTTGGCAACATGTTGATGCCAAGGCTTGTACTAGAGATGGTGCATTTTAG 117  
 Db 83 GTTTCCTAT-----ATGTTGATGCTTTGGGTTGTACTAAAGAATGGTAACTCTTG 133  
 QY 118 CTATGCATATGCCACGTTTCAGAAGGAAGTCCCAAAACCTATATGACCAATTTGTTG 177  
 Db 134 CTTTGGGATATGCCACGTTTCAGAAGGAAGTCCGACAAATCCCATATGATCAATTTGTTG 193  
 QY 178 CTCAGGCTATAAGGGTTGCAACTATTACAGTGTCTAAAGGAGATTTGATTGTGAAGGAGA 237  
 Db 194 CTCAGGCTATAAGGGTTGTAATTTATATAGTGTCTTTTCGGGAGATTTATTTGCGAAGGAGA 253  
 QY 238 ATCTGACCTTAGAACCCCAAAAGATTGTACCTTCGAATGTGTATGATACACAGATTCCTATTTC 297  
 Db 254 ATCTGACCCAAAACCCAAAAGCTTGCCTCCCTTAAATTTGTGATACAAAATATTTCCTATTTC 313  
 QY 298 AAAATGCTCTGCTTCAGAAGGAAGATGATAATTAACCCACTGGATGACCACTTGTGTTG 357  
 Db 314 AGATGTCCTCCGTTTCAGAAGGAAGAAATCGCTAATTTATCCACCGAATGTACCAATGCTG 373  
 QY 358 CACGGGCTATCAGGGTTGCTACTATTTCGATCAAGATGGTGTGATTTTGTCTGTGAAGGAGA 417  
 Db 374 CACAGGGTATAGGATTTGCTACTATTTCGATAAAATGCAAGTTTGTATGTGAAGGAGA 433  
 QY 418 GAGTCTCTGAACCCCAAGAC 435  
 Db 434 GAGTGTATGAACCCCAAGGC 451

RESULT 10  
 BI432721

LOCUS  
 DEFINITION  
 EST535482 P. infestans-challenged potato leaf, compatible reaction  
 Solanum tuberosum cDNA clone PPCAU32 5' sequence, mRNA sequence.  
 BI432721  
 BI432721.1 GI:15257411  
 EST.  
 Solanum tuberosum (potato)  
 Solanum tuberosum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamiales; Solanales; Solanaceae; Solanum.  
 1 (bases 1 to 628)  
 Restrepo,S., Griffiths,H.M., Smart,C.D., Cho,J., Chiemiango,A.,  
 Bougri,O., Buell,C.R., Ronning,C.M., Fry,W.E. and Baker,B.  
 Generation of ESTs from Potato Leaves Challenged with Phytophthora  
 infestans, Compatible Interaction  
 Unpublished (2000)  
 Contact: Robin Buell  
 The Institute for Genomic Research  
 9712 Medical Center Dr, Rockville, MD 20850, USA  
 Email: potato-array@tigr.org  
 This clone can be obtained from the University of Arizona Genomics  
 Institute. Orders can be made through URL:  
 http://genome.arizona.edu/orders/

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT



```

FEATURES
  source
    Seq primer: M13P-R.
    Location/Qualifiers
      1..628
        /organism="Solanum tuberosum"
        /mol_type="mRNA"
        /culturivar="Kennebec"
        /db_xref="taxon:4113"
        /clone="PPCAU32"
        /tissue_type="leaf"
        /dev_stage="6 week old"
        /lab_host="SOLR"
        /clone_lib="p. infestans-challenged potato leaf,
        compatible reaction"
        /notes="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
        XhoI; supplier: Cornell University, Fry lab; sequencing:
        The Institute for Genomic Research; whole plants were
        challenged with 20,000 sporangia/ml of the compatible P.
        infestans isolate US 940480. Leaf tissue was collected at
        3, 6, 9, 12, 24, 48, 72 hours after inoculation and frozen
        in liquid nitrogen immediately upon removal. Kennebec
        plants showed first symptoms of infection at 48 hours
        after inoculation. NOTE: We cannot exclude the
        possibility that this sequence is actually derived from
        Phytophthora rather than potato."

ORIGIN
  Query Match      51.2%; Score 270.8; DB 4; Length 628;
  Best Local Similarity 79.7%; Pred. No. 1.1e-63;
  Matches 349; Conservative 0; Mismatches 77; Indels 12; Gaps 2;

QY 1 CATAATGGCTGTTTCACA---AAGTTAGCTTCCTTGGCTTACTTGGTCTTGGATGGAT 57
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 26 CATAATGGCTGTTTCACAAGGAAGTTAGTTTGGTTGCTTACTACTAAATTTGTTGGAAT 85

QY 58 GTTCTTACTTCGGAACATGTTTCATCCCAAGGCTTGTACTAGAGAATGGTCAATTTAG 117
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 86 GTTCTTAT-----ATGTTGATGCTTTGGGTTGTACTAAAGAATGGTAACTTGG 136

QY 118 CTATGGCATATGCCACGTTTCAGAAGGAAGTCCCAAAACCTTATATGACCAATTTGTTG 177
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 137 CTTTGGGATATGCCACGTTTCAGAAGGAAGTCCGACAAATCCCATATGATCAATTTGTTG 196

QY 178 CTCAGGCTATAGGGTTGCAACTATTACAGTGTCTAAAGGAGATTTGATTTGGAAGGAGA 237
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 197 CTCAGGCTATAGGGTTGTAATTTATTATAGTGTCTTTCCGGAGATTTTATTTCGAAGGAGA 256

QY 238 ATCTGACCTTAGAACCCCAAGATTTGTACCTTCGAATGTGTACACAGATTTGCTTATTC 297
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 257 ATCTGACCCCAAAACCCCAAGAGCTTGCCCCCTTAAATTTGTGATACAAATTTGCCTTATTC 316

QY 298 AAAATGTCCTCGTTTCAGAAGGAAGATGATAATTAACCCACCTGGATGCACCACTTTGTTG 357
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 317 AGAGTGTCCCGTTTCAGAGGAATAATCGCTAATTTATCCACCGGATGTACCATGCTG 376

QY 358 CACGGGCTATCAGGGTTGCTACTATTTCGATCAAGATGGTGATTTTGTCTGGAAGGAGA 417
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 377 CACAGGGTATAGGAATTTGCTACTATTTCGATAAAATGGCAAGTTGTTGTATGGAAGGAGA 436

QY 418 GAGTCCTGACCCCAAGAC 435
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 437 GAGTGAATGAACCCCAAGGC 454

RESULT 11
BF153853
LOCUS
DEFINITION
  BF153853 672 bp mRNA linear EST 11-SEP-2002
  032C04 Mature tuber lambda ZAP Solanum tuberosum cDNA 5' similar to
  PROTEINASE INHIBITOR TYPE II CM7 PRECURSOR . . . sp|Q43652, mRNA
  sequence.
ACCESSION
  BF153853
VERSION
  BF153853.1 GI:11035793
KEYWORDS
  EST.
  Solanum tuberosum (potato)

ORGANISM
  Solanum tuberosum
  Sukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE
  1 (bases 1 to 672)
AUTHORS
  Crookshanks,M., Emmersen,J., Welinder,K.G. and Nielsen,K.L.
TITLE
  The potato tuber transcriptome: analysis of 6077 expressed sequence
  tags
JOURNAL
  FEBS Lett. 506 (2), 123-126 (2001)
MEDLINE
  21475600
PUBMED
  11591384
COMMENT
  Contact: Karen G. Welinder
  Institut for bioteknologi
  Aalborg Universitet
  Sohngaardsholmsvej 49, 9000 Aalborg, Denmark
  Tel: +45 96358467
  Fax: +45 98141808
  Email: kgw@bio.auc.dk
  High quality sequence stop: 672
  POLYA=yes.
FEATURES
  source
    Location/Qualifiers
      1..672
        /organism="Solanum tuberosum"
        /mol_type="mRNA"
        /culturivar="Field grown Kurag"
        /db_xref="taxon:4113"
        /tissue_type="tuber"
        /clone_lib="Mature tuber lambda ZAP"
        /note="Vector: Lambda ZAP"

ORIGIN
  Query Match      50.9%; Score 269.4; DB 2; Length 672;
  Best Local Similarity 78.3%; Pred. No. 2.8e-63;
  Matches 353; Conservative 0; Mismatches 86; Indels 12; Gaps 2;

QY 1 CATATGGCTGTTTCACA---AAGTTAGCTTCCTTGGCTTACTTGGTCTTGGATGGAT 57
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 28 CATCATGGCTGTTTCACAGGAAGTTAAATTTGTTGCTTACCTACTAAATTTGTTCTTGAAT 87

QY 58 GTTCTTACTTGT-----CGAAACATGTTGATGCCAAGGCTTGTACTAGAGAATGTGG 108
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 88 ATTTCTACTTGTAGCGTGTGGACATGTTGATGCGAAGGCTTGTACTTTAGAATGTGG 147

QY 109 TCATTTTACTATGGCATATGCCACGTTTCAGAGGAAGTCCCAAAACCTTATATGCAC 168
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 148 TAACTCTGGATTTGGGATATGCCACGTTTAGAAGGAAGTCCGACAAATCCCATATGCAT 207

QY 169 CAATTGTTCTCAGGCTATAAGGTTGCAACTATTACAGTGTCTAAAGGAGATTTGATTG 228
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 208 CAATTGTTCTCAGGCTATAAGGTTGTAAATTTATTATAGTGTCTTTTGGGAGATTTATTG 267

QY 229 TGAAGGAGAAATCTGACCCCTAGAAACCCAAAGAAATTTGTACCTTCGAATGTGATACACAGAT 288
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 268 TGAAGGAGAAATCTGACCCCAAAACCCAAAGAAATTTGTACCTTCGAATGTGATACAAATAT 327

QY 289 TGCCTATTCAAAATGTCTCGTTTCAGAGGAAGATGATAATTAACCCACTCGGATGCAC 348
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 328 TGCCTATTCAAGATGTTCCTCGTTTCAAAAGGAAACAACTAAATTTATCCACCGGATGTAC 387

QY 349 CACTTGTTCACGGGCTATCAGGGTTGCTACTATTTCGATCAAGATGGTGTGATTTCCTG 408
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 388 CACATGCTGCACAGGGGTACAAGGGTTGCTACTATTTCGTTAAATAATGGCAAGTTTGTATG 447

QY 409 TGAAGGAGAGAGTCTCTGAACCCCAAGACCACT 439
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 448 TGAAGGAGAGAGTGTATGAGCCCCCAAGCAAAAT 478

RESULT 12
BI436259
LOCUS
DEFINITION
  BI436259 675 bp mRNA linear EST 10-MAR-2003
  EST539020 cSTE Solanum tuberosum cDNA clone cSTE22115 5' sequence,
  mRNA sequence.
```

```

ACCESSION BI436259
VERSION BI436259.1 GI:15260949
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 675)
AUTHORS van der Hoeven,R., Bezzerides,J., Bachem,C., Visser,R., Cho,J.,
Chieningo,A., Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and
Baker,B.
TITLE Generation of ESTs from in vitro grown microtubers
JOURNAL Unpublished (2001)
COMMENT Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: M13P-R.
FEATURES             Location/Qualifiers
     source            1..675
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Bintje"
                     /db_xref="taxon:4113"
                     /clone="cSTRE22115"
                     /tissue_type="axillary buds of stem explants; growing
                     sink-tubers"
                     /dev_stage="7, 8 and 10 days"
                     /lab_host="SOLR"
                     /clone_lib="cSTg"
                     /notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; Tissue supplied by Christian Bachem and Richard
                     Visser (Department of Plant Breeding, Wageningen
                     University, The Netherlands). The cSTA libraries will
                     attempt to capture the induction and initiation/initial
                     growth of the tuber in an in vitro system as described in
                     Bachem et al. (Plant Journal, 1996). Small microtubers
                     develop from axillary buds attached to stem explants when
                     placed on a high sucrose medium (10%). Visible
                     morphological changes occur synchronously at day five in
                     the axillary buds. The first library, cSTA (1-20) consists
                     of axillary buds harvested on days 1-3. This targets
                     those genes involved in induction of the microtubers. The
                     following libraries, cSTA (21-40) and cSTA (41-60),
                     capture genes involved in tuber initiation and outgrowth.
                     This library is noted as P3 in Tanksley lab notebooks."
ORIGIN
Query Match      50.9%; Score 269.4; DB 4; Length 675;
Best Local Similarity 78.9%; Pred. No. 2.8e-63;
Matches 337; Conservative 0; Mismatches 81; Indels 9; Gaps 1;

QY 18  AAGTAGCTTCCTGCTTGGCTTACTTCTTCTCGATGGATGTTCTACTTT-----G 68
DB 28  AAGTTAGTTTCCTTGCTTACTTACTTCTTCTTCTTGGATTATGCTTGTGAAGCAGG 87

QY 69  CGAAACATGTTGATGCCAAGCTTGCTACTAGAGAAATGGTCTATTTTAGCTATGGCATAT 128
DB 88  TGGACATGTTGATGCCAAGCTTGCTACTTAAAGAAATGGTGAATCTTGGCTATGGGATAT 147

QY 129 GCCACGTTCAGAGGAAGTCCCAAAAACCTATATGACCAATTTGTTGCTCAGGCTATA 188
DB 148 GCCACGTTCAGAGGAAGTACGGAAATCCCATATGACCAACTGTTGTGCAAGGTATA 207

QY 189 AGGGTTGCACTATTACAGTGCCTAAAGGAGATTTGATTTCTGAAGGAGATCTGACCTTA 248
DB 208 AAGTTGCAATATTATTAGTGCCTAAATGGGATTTTCATTTGCGAAGGACAACTGACCCAA 267

QY 249 GAAACCCAAAAGATTGTACCTTCGAATGTGATACACAGATTGCTTATTCAAATGTCCTC 308

```

---

```

Db 268 AAAACCCAAAAGCTTCCCGGAAATTTGATGCACACATTTGCTATTCAAAGTGTCTTC 327
QY 309 GTTTCAGAGGAAGATGATAATTAAACCCACCTGGATGCACCACTTGTTCACCGGGTATC 368
Db 328 GTTCCAGAGGAAGAGAGCTAATATATCCACGGATGTACCAATGTTGACAGAGGTACA 387
QY 369 AGGGTTGCTACTATTTTCGATCAAGATGGTGATTTTCTGTGTAAGGAGAGAGTCTTGAAC 428
Db 388 AGGATTGCTACTATTTTGGTAAAGACGGCAAGTTTGTGTGAAGGAGAGAGTATTGAAC 447
QY 429 CCAAGAC 435
Db 448 CCAAGGC 454

RESULT 13
BF153879
LOCUS
DEFINITION
BFI53879.1 GI:11035819
ACCESSION
VERSION BFI53879
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 677)
AUTHORS Crookshanks,M., Emmersen,J., Welinder,K.G. and Nielsen,K.L.
TITLE The potato tuber transcriptome: analysis of 6077 expressed sequence
tags
JOURNAL FEBS Lett. 506 (2), 123-126 (2001)
MEDLINE 21475600
PUBMED 11591384
COMMENT Contact: Karen G. Welinder
Institut for bioteknologi
Aalborg Universitet
Sohngaardsholmsvej 49, 9000 Aalborg, Denmark
Tel: +45 96355467
Fax: +45 98141808
Email: kgw@bio.auc.dk
High quality sequence stop: 677
POLYA=yes.
FEATURES             Location/Qualifiers
     source            1..677
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Field grown Kurat"
                     /db_xref="taxon:4113"
                     /tissue_type="tuber"
                     /clone_lib="Mature tuber lambda ZAP"
                     /notes="Vector: Lambda ZAP"
ORIGIN
Query Match      50.9%; Score 269.4; DB 2; Length 677;
Best Local Similarity 78.3%; Pred. No. 2.8e-63;
Matches 353; Conservative 0; Mismatches 86; Indels 12; Gaps 2;

QY 1  CATAATGGCTGTTTCAACA-----AAGTTAGCTTCCCTTGGCTTGTCTTCTTCTTGGATGGAT 57
DB 28  CATCATGGCTGTTTCAAGAGAGTTAATTTGGTTGCTTACCTACTAATTTGTTCTTGAAT 87

QY 58  GTTTCCTACTTG-----CGAAACATGTTGATGCCAAGGCTTGTGTAGAGAAATGTGG 108
DB 88  ATTCTACTTGTGTAGCGTGTGGAAACATGTTGATGCGAAGGCTTGTACTTTTAGAATGTGG 147

QY 109 TCATTTTACTAGTATGGCATATGCCACGTTTCAGAAGGAAGTCCCAAAAACCTATATGACAC 168
DB 148 TAATCTTGGATTGGGATATGCCCGCTTTAGAAGGAAGTCCGACAAATCCCATATGCAAT 207

```

```

QY 169 CAATTGTTGCTCAGGCTATAGGGTTGCAACTATTACAGTGTAAAGGAGATTGATTTG 228
DB 208 CAATGTTGCTCAGGCTATAGGGTTGTAATATTATAGTGTCTTTGGGAGATTATTTG 267
QY 229 TGAAGAGAAATCTGACCTAGAACCCCAAAAGATTGTACCTTCGAAATGTGATACACAGAT 288
DB 268 TGAAGAGAAATCTGACCCAAAACCCAAAGCTTGCCCTTAAATTTGTGATACAAATAT 327
QY 289 TGCTTATTCAAAATGTCCTCGTTCAGAAAGGAAGATGATTAATTAACCCCACTGGATGCAC 348
DB 328 TGCTTATTCAAAGATGTTCCCGTTCAAAAGGAAAAACACTAATTTATCCCAACCGGATGAC 387
QY 349 CACTTGTTCACGGCTATCAGGGTTGCTACTATTTCGATCAAGATGGTGTATTGTCG 408
DB 388 CACATGCTGCACAGGGTACAGGGTTGCTACTATTTTCGGTAAAAATGGCAAGTTTGTATG 447
QY 409 TGAAGAGAGAGTCTGAAACCCCAAGACCACT 439
DB 448 TGAAGAGAGAGTATGATGAGCCCAAGGCAAT 478

RESULT 14
BF153508 591 bp mRNA linear EST 11-SEP-2002
LOCUS 027E11 Mature tuber lambda ZAP Solanum tuberosum cDNA 5' similar to
DEFINITION PROTEINASE INHIBITOR TYPE II CM7 PRECURSOR . . . sp|Q43652. mRNA
sequence.
ACCESSION BF153508
VERSION BF153508.1 GI:11035448
KEYWORDS Solanum tuberosum (potato)
SOURCE Solanum tuberosum
ORGANISM Solanum tuberosum (potato)

REFERENCE 1 (bases 1 to 591)
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
TITLE The potato tuber transcriptome: analysis of 6077 expressed sequence
tags
JOURNAL FEBS Lett. 506 (2), 123-126 (2001)
MEDLINE 21475600
PUBMED 11591384
COMMENT Contact: Karen G. Wellinder
Institut for bioteknologi
Aalborg Universitet
Sohngaardsholmsvej 49, 9000 Aalborg, Denmark
Tel: +45 96358467
Fax: +45 98141808
Email: kgw@bio.auc.dk
High quality sequence stop: 591
POLYA=No.

FEATURES
source
1. .591
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Field grown Kuras"
/db_xref="taxon:4113"
/tissue_type="Tuber"
/clone_lib="Mature tuber lambda ZAP"
/notes="Vector: Lambda ZAP"

ORIGIN
Query Match 50.4%; Score 267.8; DB 2; Length 591;
Best Local Similarity 78.0%; Pred. No. 7.5e-63;
Matches 352; Conservative 0; Mismatches 87; Indels 12; Gaps 2;

QY 1 CATAATGGCTGTTACACA---AAGTAGCTTCCTGCTGCTACTGTTCTTCGATGGAT 57
DB 9 CATCATGGCTGTTACAAAGGAAGTAAATTCGTTGCTTACTACTAATTCGTTCTTCTGAAAT 68

QY 58 GTTCTTACTTGT-----CGAAACATGTTGATGCCAGGCTGTGACTAGAAATGTGG 108
DB 69 ATTTCTACTTGTGTAGCGGTGGAAACATGTTGATGCGAGGCTGTGACTTTAGAAATGTGG 128

```

```

QY 109 TCATTTTAGCTATGCCATATGCCACGTTTCAGAAAGAGTCCCAAAAACCTATATGCAC 168
DB 129 TAATCTTGGATTTGGGATATGCCACGTTTCAGAAAGAGTCCCAAAAACCTATATGCAT 188
QY 169 CAATTGTTGCTCAGGCTATAGGGTTGCAACTATTACAGTGTAAAGGAGATTGATTTG 228
DB 189 CAATTGTTGCTCAGGCTATAGGGTTGTAATTTATATAGTGTCTTTGGGAGATTATTTG 248
QY 229 TGAAGAGAAATCTGACCTAGAACCCCAAAAGATTGTACCTTCGAAATGTGATACACAGAT 288
DB 249 TGAAGAGAAATCTGACCCAAAACCCAAAGCTTGCCCTTAAATTTGTGATACAAATAT 308
QY 289 TGCTTATTCAAAATGTCCTCGTTCAGAAAGGAAGATGATTAATTAACCCCACTGGATGCAC 348
DB 309 TGCTTATTCAAAGATGTTCCCGTTCAAAAGGAAAAACACTAATTTATTCCTCCACCGGATGAC 368
QY 349 CACTTGTTCACGGCTATCAGGGTTGCTACTATTTCGATCAAGATGGTGTATTGTCGTG 408
DB 369 CACATGCTGCACAGGGTACAGGGTTGCTACTATTTCGGTAAAAATGGCAAGTTTGTATG 428
QY 409 TGAAGAGAGAGTCTGAAACCCCAAGACCACT 439
DB 429 TGAAGAGAGAGTATGATGAGCCCAAGGCAAT 459

RESULT 15
BF153508 779 bp mRNA linear EST 16-AUG-2001
LOCUS EST531828 tomato callus, TAMU Lycopersicon esculentum cDNA clone
DEFINITION GLEC6619 5' end, mRNA sequence.
ACCESSION BF153508
VERSION BF153508.1 GI:15194442
KEYWORDS Lycopersicon esculentum (tomato)
SOURCE Lycopersicon esculentum
ORGANISM Lycopersicon esculentum (tomato)

REFERENCE 1 (bases 1 to 779)
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
TITLE Generation of ESTs from tomato callus tissue
JOURNAL Unpublished (1999)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.

FEATURES
source
1. .779
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="GLEC6619"
/tissue_type="callus"
/dev_stage="25-40 days old"
/lab_host="XLI-Blue MRP"
/clone_lib="tomato callus, TAMU"
/notes="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Giovannoni laboratory; cLEC - Cotyledons
of seedlings 7-10 days post-germination were excised, cut
at both ends and placed on MS medium with no selection.
Mixed callus was harvested at 25 and 40 days and included
undifferentiated masses. Tomato Callus EST Library"

ORIGIN
Query Match 50.5%; Score 267.2; DB 4; Length 779;
Best Local Similarity 82.9%; Pred. No. 1.2e-62;
Matches 305; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

```

Qy	68	CGCAAA	CAATGTTG	ATGCGCA	AGGCTTG	TACTAG	AGAAATG	TGGTCA	TTTTAG	CTATG	CGCATA	127
Db	35	GTGGA	CAATGTTG	ATGCGCA	AGGCTTG	TACTAG	AGAAATG	TGGTCA	TTTTAG	CTATG	CGGATA	94
Qy	128	TGCCA	CGCTTC	GAGAGG	AAAGTCC	CCAAAA	CCATATG	CACCAAT	TGTTGC	TCAAGC	CTAT	187
Db	95	TGCCA	CGCTTC	CAAGGA	AGTCCG	CAGAAT	CCCATATG	CACCAAT	TGTTG	TTCAGG	TTTT	154
Qy	188	AAGGT	TTCCAA	CTATTAC	AGTGCT	AAAGG	AGATTTG	ATTGTA	AGGAGAA	TCTGAC	CCCT	247
Db	155	AAAGT	TTCCAA	TTATTA	TAGTGT	TAATGGG	ACTTTT	ATTGTA	AGGACAA	TCTGAC	CCCA	214
Qy	248	AGAAA	CCCAA	AGATG	TGACCTT	CGAATG	TGTGATA	CACAGAT	TGCTTAT	TCAAAAT	TGTCCT	307
Db	215	AGAAA	CCCAA	ACCTTG	CGCCCTT	AAATGT	GTGATCC	ACATAT	TGCTTAT	TCAAAAT	TGTCCT	274
Qy	308	CGTTC	GAGAGG	AAAGAT	GATTAAT	TAAACCA	CTGGATG	CACCACT	TGTTG	TGCACG	GGCTAT	367
Db	275	CGTTC	GAGAGG	AAAGAC	GTTTAT	TATCCCA	CTGGATG	TACCA	CGTGTG	TGCACG	GGATAC	334
Qy	368	CAGGT	TTGCTAC	TATTTCC	ATCGAT	TGTTGAT	TTTGTCTG	TGTAAGG	AGAGAGT	CTCTGA	A	427
Db	335	AAAGT	TTGCTAC	TATTTCCG	TAAAGAT	TGCAAGT	TTTGTG	TGTAAGG	AGAGAT	TGATGA	A	394
Qy	428	CCC	AAAGAC	435								
Db	395	CCC	AAAGC	402								

Search completed: January 22, 2005, 07:16:48  
Job time : 2589 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2005, 05:24:58 ; Search time 439 Seconds  
(without alignments)  
6923.861 Million cell updates/sec

Title: US-10-725-829-1  
Perfect score: 529  
Sequence: 1 cataatggctgttcacaag.....ctgtaatggactttattg 529

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0 .

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Databases : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCT05\_PUBCOMB.seq\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq\*
- 18: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq\*
- 19: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq\*
- 20: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq\*
- 21: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	529	100.0	529	18	US-10-725-829-1 Sequence 1, Appli
2	288.4	54.5	692	18	US-10-725-829-3 Sequence 3, Appli
3	178.8	33.8	1360	14	US-09-812-502-2 Sequence 2, Appli
4	178.8	33.8	1360	14	US-10-164-961-2 Sequence 2, Appli
5	178.8	33.8	1360	15	US-10-157-622-2 Sequence 2, Appli
6	149	28.2	1104	14	US-09-812-502-1 Sequence 1, Appli
7	149	28.2	1104	14	US-10-164-961-1 Sequence 1, Appli
8	149	28.2	1104	15	US-10-157-622-1 Sequence 1, Appli
9	149	28.2	1104	15	US-10-072-809A-56 Sequence 56, Appli
10	35.6	6.7	680	13	US-10-027-632-281276 Sequence 281276,
11	35.6	6.7	680	15	US-10-027-632-281276 Sequence 281276,
12	35.6	6.7	108359	14	US-10-191-807-3 Sequence 3, Appli

c	13	35.2	6.7	6134	16	US-10-221-714A-285	Sequence 285, App
c	14	35.2	6.7	6650	15	US-10-311-455-611	Sequence 611, App
c	15	35.2	6.7	10433	15	US-10-311-455-352	Sequence 352, App
	16	35	6.6	9964	15	US-10-311-455-71	Sequence 71, Appl
	17	34.8	6.6	632	13	US-10-027-632-96703	Sequence 96703, A
	18	34.8	6.6	632	15	US-10-027-632-96703	Sequence 96703, A
	19	34.8	6.6	6866	15	US-10-204-708-19	Sequence 19, Appl
	20	34.8	6.6	6866	15	US-10-311-455-639	Sequence 639, App
	21	34.8	6.6	152330	13	US-10-087-192-1834	Sequence 1834, App
c	22	34.6	6.5	346	9	US-09-880-107-860	Sequence 860, App
c	23	34.4	6.5	789	15	US-10-369-493-42185	Sequence 42185, A
c	24	34.4	6.5	63155	15	US-10-292-798-449	Sequence 449, App
c	25	34.4	6.5	193303	15	US-10-081-327-37	Sequence 37, Appl
c	26	34.4	6.5	193303	15	US-10-081-327-44	Sequence 44, Appl
c	27	34.2	6.5	1262	9	US-09-070-927A-759	Sequence 759, App
c	28	34.2	6.5	1479	16	US-10-425-115-36514	Sequence 36514, A
c	29	34.2	6.5	1479	18	US-10-425-115-22922	Sequence 22922, A
c	30	34	6.4	32249	10	US-09-764-891-7477	Sequence 7477, Ap
c	31	34	6.4	75899	9	US-09-854-883-243	Sequence 243, App
c	32	34	6.4	75899	15	US-10-360-510-243	Sequence 243, App
c	33	34	6.4	518360	17	US-10-367-094-123	Sequence 125, App
c	34	33.8	6.4	1059	13	US-10-027-632-250275	Sequence 250275,
c	35	33.8	6.4	1059	15	US-10-027-632-250275	Sequence 250275,
c	36	33.8	6.4	5473	15	US-10-037-270-260	Sequence 260, App
c	37	33.8	6.4	5473	15	US-10-117-722-260	Sequence 260, App
c	38	33.8	6.4	7545	17	US-10-473-575-25	Sequence 25, Appl
c	39	33.4	6.3	255	10	US-09-930-213-697	Sequence 697, App
c	40	33.4	6.3	648	9	US-09-861-451A-25	Sequence 25, Appl
c	41	33.4	6.3	668	13	US-10-027-632-144414	Sequence 144414,
c	42	33.4	6.3	668	15	US-10-027-632-144414	Sequence 144414,
c	43	33.4	6.3	2447	16	US-10-264-237-715	Sequence 715, App
c	44	33.4	6.3	39560	15	US-10-085-117-118	Sequence 118, App
c	45	33.4	6.3	96597	11	US-09-997-722-289	Sequence 289, App

ALIGNMENTS

RESULT 1

US-10-725-829-1  
; Sequence 1, Application US/10725829  
; Publication No. US20040205846A1  
; GENERAL INFORMATION:  
; APPLICANT: Chye, Mee Len  
; APPLICANT: Xu, Zeng-Fu  
; APPLICANT: Sin, Suk Fong  
; TITLE OF INVENTION: Genetically Modified Plants with Heterogenous Proteinase Inhibitor  
; TITLE OF INVENTION: SapIN2A or SapIN2B and Methods of Use Thereof for the Inhibition  
; TITLE OF INVENTION: Of Trypsin- and Chymotrypsin-Like Activities  
; FILE REFERENCE: 9661-043-999  
; CURRENT APPLICATION NUMBER: US/10/725,829  
; CURRENT FILING DATE: 2003-12-01  
; PRIOR APPLICATION NUMBER: 60/429,992  
; PRIOR FILING DATE: 2002-11-29  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1  
; LENGTH: 529  
; TYPE: DNA  
; ORGANISM: Solanum americanum  
US-10-725-829-1

Query Match	100.0%	Score 529;	DB 18;	Length 529;
Best Local Similarity	100.0%;	Pred. No. 3.8e-142;		
Matches 529;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	CATAATGGCTGTTCACAAAGTTAGCTTCCTTGGCTGCTACTTGTCTTGGATGGATGTT	60	
Db	1	CATAATGGCTGTTCACAAAGTTAGCTTCCTTGGCTGCTACTTGTCTTGGATGGATGTT	60	
Qy	61	TCTACTTGGCAACATGTTGATGCCAAGGCTTCTACTAGAGATGTTGTCATTTCAGCTA	120	
Db	61	TCTACTTGGCAACATGTTGATGCCAAGGCTTCTACTAGAGATGTTGTCATTTCAGCTA	120	

```
QY 121 TGGCATATGCCCGTTGCAAGGAAGTCCCAAAACCTATATGACCAATTTGTTGCTC 180
Db 121 TGGCATATGCCCGTTGCAAGGAAGTCCCAAAACCTATATGACCAATTTGTTGCTC 180
QY 181 AGGCTATAAGGGTTGCAACTATTACAGTGTGTAAGGAGATTTGATTGTAAGGAGAATC 240
Db 181 AGGCTATAAGGGTTGCAACTATTACAGTGTGTAAGGAGATTTGATTGTAAGGAGAATC 240
QY 241 TGAACCTAGAACCCAAAGATTTGACCTTCGATGTGATACACAGATTTCTATTCAAA 300
Db 241 TGAACCTAGAACCCAAAGATTTGACCTTCGATGTGATACACAGATTTCTATTCAAA 300
QY 301 ATGCTCTCGTTTCAAGAGGAAGATGATAATTAACCCACCTGGATGACCACTTTGTTGCAC 360
Db 301 ATGCTCTCGTTTCAAGAGGAAGATGATAATTAACCCACCTGGATGACCACTTTGTTGCAC 360
QY 361 GGGCTATCAGGGTTGCTACTATTTCGATCAAGATGGTGAATTTTGTCTGTGAAGGAGAG 420
Db 361 GGGCTATCAGGGTTGCTACTATTTCGATCAAGATGGTGAATTTTGTCTGTGAAGGAGAG 420
QY 421 TCCGTGAACCAAGACCACTGCTTATTTCATCAATCAATGATGTTGTTATCTATCAAAAA 480
Db 421 TCCGTGAACCAAGACCACTGCTTATTTCATCAATCAATGATGTTGTTATCTATCAAAAA 480
QY 481 AAATATGTATGCATGATATATGCTGGTTACTGTAATGTGACITTTATTG 529
Db 481 AAATATGTATGCATGATATATGCTGGTTACTGTAATGTGACITTTATTG 529

RESULT 2
US-10-725-829-3
; Sequence 3, Application US/10725829
; Publication No. US20040205846A1
; GENERAL INFORMATION:
; APPLICANT: Chye, Mee Len
; APPLICANT: Xu, Zeng-Fu
; APPLICANT: Sin, Suk Fong
; TITLE OF INVENTION: Genetically Modified Plants with Heterogenous Proteinase Inhibitor
; TITLE OF INVENTION: SAPIZA or SAPIZB and Methods of Use Thereof for the Inhibition
; FILE REFERENCE: 9661-043-999
; CURRENT APPLICATION NUMBER: US/10/725,829
; CURRENT FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: 60/429,992
; PRIOR FILING DATE: 2002-11-29
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 692
; TYPE: DNA
; ORGANISM: Solanum americanum
US-10-725-829-3

Query Match 54.5%; Score 288.4; DB 18; Length 692;
Best Local Similarity 80.5%; Pred. No. 1.3e-72;
Matches 352; Conservative 0; Mismatches 76; Indels 9; Gaps 1;

QY 18 AGATTAGCTTCCTTGTGCTTACTTGTTCCTCGATGGATGTTTCTACTTTGCG----- 70
Db 49 AAGTTAGTTCCCTTGTCTTACTTACTTGTTCCTCGATGATGTTTCTTACATGTAAGCGCG 108
QY 71 --AAACATGTTGATGTCNAGGCTTGCTAGAGAAATGTGTCATTTTAGTATGGCATAT 128
Db 109 TAAAAACATGTTGATGTCNAGGCTTGCTAGAGAAATGTGTCATTTTAGTATGGCATAT 168
QY 129 GCCCAGCTTCAGAGGAAGTCCCAAAACCTATATGACCAATTTGTTGCTCAGGCTATA 188
Db 169 GCCCGGTTTCAGAGGAAGTCCCGAAATCCCATATGACGAATTTGTTGCTCAGGCTATA 228
QY 189 AGGGTTGCAACTATTACAGTGTGTAAGGAGATTTGATTGTGAAGGAGAAATCTGACCTTA 248
Db 229 AAGGTTGCAACTATTATAGTGTCTAAATGGGACTTTTATTTTCGGAAGGAAGTTCTTGACCTTA 288
```

```
QY 249 GAAACCCAAAAGATTTGATCCTTCGAATGTGATACACAGATTTGCTTATTCAAATGTCTC 308
Db 289 AAAACCCAAAATACTTCCCTCTTATTGTTGATGGAGATATTGCTATTCAAATGTCTCC 348
QY 309 GTTCAGAGGAAGATGATAATTAACCCACCTGGATGACCACTTTGTGACCGGGCTATC 368
Db 349 GTTCAGAGGAAGATGATAATTAATATATCCACCGGATGACCACTCTGTGACCGGGTACA 408
QY 369 AGGGTTGCTACTATTTCGATCAAGATGTGATTTTCTGTGAAGGAGAGAGTCTGAAC 428
Db 409 AGGGTTGCTACTATTTTAGTAAAGAAAGGTGAGTTTGTGTGAAGGAGAGATGATGAAC 468
QY 429 CCAAGACCACTGCTTAT 445
Db 469 CCAACGTTATTCTAAT 485

RESULT 3
US-09-812-502-2
; Sequence 2, Application US/09812502
; Publication No. US20030027303A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; TITLE OF INVENTION: SEQUENCES ENCODING SAME
; FILE REFERENCE: 9748B
; CURRENT APPLICATION NUMBER: US/09/812,502
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US/09/431,500
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 08/454,295
; PRIOR FILING DATE: 1995-09-01
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1360
; TYPE: DNA
; ORGANISM: Nicotiana glauca
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (97)..(1200)
US-09-812-502-2

Query Match 33.8%; Score 178.8; DB 10; Length 1360;
Best Local Similarity 67.3%; Pred. No. 7.9e-41;
Matches 303; Conservative 0; Mismatches 132; Indels 15; Gaps 3;

QY 5 ATGGCTGTTCAAAAGTTAGCTTCCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTA 64
Db 10 ATGGCTGTTCAAGATTTAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 69
QY 65 CT-----TGCAGAAACATGTTGATGCAAGGCTTGTACTAGAGAAATGTGTCATTT---TT 115
Db 70 GTAAGCAATGTGGAACATGCGATGCAAGGCTTGTACCTTAACCTGTGATCCAGAAAT 129
QY 116 AGCTATGGCATATGCCACGTTTCAGAGGAAGTCCCAAAACCTATATGACCAATTTGT 175
Db 130 GCCTATGGAGTTTGCCTCGGTTTCAGAGGAAGAAAGAAATGATCGGATATGACCAACTGT 189
QY 176 TGCTCAGGCTATAGGGTTGCAACTATTACAGTGTCTTAAGGAGATTTGATTTGTGAAGGA 235
Db 190 TCGCAGGCAACGAGGGTTGTAGTACTTTCAGTGTGATGGAAGTCTTCTTCTTCTTCTTCT 249
QY 236 GAATCTGACCTAGAAACCAAAAGATTTGACCTTCCAAATGTGATACACAGATTTGCTTAT 295
Db 250 GAGTCTGATCTTAGAAATCCAAAGGCTTGTACCTTAACTGATCCAGAAATTTGCTTAT 309
QY 296 TCAAAATGTCCTCGTTTCAGAGGAAGATGATAATTTAAACCCACTCGATGACCACTTGT 355
Db 296 TCAAAATGTCCTCGTTTCAGAGGAAGATGATAATTTAAACCCACTCGATGACCACTTGT 355
```

```
Db 310 GGAGTTTCCCGCTTCAGAAAGAAAGAG-----AATGATCGGATATGCACCAACTGT 363
Qy 356 TGCACGGGCTATCAGGTTCTACTATTTCGATCAAGATGGTGAATTTGTCTGTGAAGGA 415
Db 364 TGGCAGGACGAAAGGTTTGAAGTACTTCAGTGATGATGGAACCTTTTGTGTGAAGGA 423
Qy 416 GAGAGTCTCAACCCCAAGCACTGCTTAT 445
Db 424 GAGTCTGATCCTAGAAATCCAAAGGCTTGT 453

RESULT 4
US-10-164-961-2
; Sequence 2, Application US/10164961
; Publication No. US20030096388A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
;               Atkinson, Angela H.
;               Heath, Robyn L.
;               Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/164,961
; FILING DATE: 07-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,295
; FILING DATE: 01-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1360 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 97..1200
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-164-961-2

Query Match 33.8%; Score 178.8; DB 14; Length 1360;
Best Local Similarity 67.3%; Pred. No. 7.9e-41;
Matches 303; Conservative 0; Mismatches 132; Indels 15; Gaps 3;

Qy 5 ATGGCTGTTCAAAAGTACTTCCTTGTGCTACTTGTTCGATGGAATGTTTCTA 64
Db 10 ATGGCTGTTCAAGAGTACTTCCTTGTGCTACTTGTTCGATGGAATGTTTCTGCTT 69
Qy 65 CT-----TGCAGAAATGTTGATGCAAGGCTTGTACTAGAAATGTTGTCATT---TT 115
Db 70 GTAAGCAATGTGGAACATGCAGATGCCAAGGCTTGTACCTTAACTTAACTGTGATCCAAGAA 129
```

## RESULT 5

```
US-10-157-622-2
; Sequence 2, Application US/10157622
; Publication No. US20030129720A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
;               Atkinson, Angela H.
;               Heath, Robyn L.
;               Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/157,622
; FILING DATE: 29-May-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,295
; FILING DATE: 01-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1360 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
```

```
Qy 116 AGCTATGGCATATGCCACAGTTTCAGAAAGAAAGTCCCCAAAAACCTATATATGCACCAATTGT 175
Db 130 GCCTATGGAGTTTGGCCCGGTTTCAGAAAGAAAGAAAGATGATCGGATATATGCACCAACTGT 189
Qy 176 TGCTCAGGCTATAAAGGTTTCAACTATTACAGTGTCTAAAGGAGATTGATTTGTGAAGGA 235
Db 190 TGGCAGGACGAAAGGTTTGAAGTACTTCAGTGATGATGGAACCTTTTGTGTGAAGGA 249
Qy 236 GAATCTGACCTAGAAACCCAAAGATTGTACCTTCGAATGTGATACACAGATTGCTTAT 295
Db 250 GAGTCTGATCCTAGAAATCCAAAGGCTTGTACCTTAAACTGTGATCCAAGAAATTCGCTAT 309
Qy 296 TCAAAATGTCCTGTTTCAGAAAGAAAGATGATTAATTAACCCACTGATGCACCACTGT 355
Db 310 GGAGTTTGGCCCGGTTTCAGAAAGAAAGAAAG-----AATGATCGGATATGCACCAACTGT 363
Qy 356 TGCACGGGCTATCAGGTTTGTCTACTATTTCGATCAAGATGGTGAATTTGTCTGTGAAGGA 415
Db 364 TGGCAGGACGAAAGGTTTGAAGTACTTCAGTGATGATGGAACCTTTTGTGTGAAGGA 423
Qy 416 GAGAGTCTCAACCCCAAGCACTGCTTAT 445
Db 424 GAGTCTGATCCTAGAAATCCAAAGGCTTGT 453
```





US-10-164-961-1

Query Match 28.2%; Score 149; DB 14; Length 1104;  
Best Local Similarity 67.2%; Pred. No. 2.9e-32;  
Matches 244; Conservative 0; Mismatches 110; Indels 9; Gaps 2;

QY 86 AAGCGTTGTACTAGAGAAATGTGTC---ATTTAGCTATGGCATATGCCACCGTTCAGAA 142  
DB 1 AAGCGTTGTACTTAACTGTGATCCAGAAATTCCTATGGAGTTTGCCTCGGCGTTCAGAA 60  
QY 143 GGAAGTCCCCAAAACCTATATGCACCAATTTGTTGCTCAGGCTATTAAGGGTTGCAACTAT 202  
DB 61 GAAAGAGAGATGATCGGATATGCACCACTGTTGCCAGGCACGAGGGTTGTAAGTAC 120  
QY 203 TACAGTGTCTAAGAGAGATTTGTTGTAAGAGAGATCTGACCTAGAAAACCCAAAGAT 262  
DB 121 TACAGTGTATGAACTTTGTTGTAAGAGAGATCTGATCCTAGAAAATCCAAAGGCT 180  
QY 263 TGTACCTTCGAATGTATACACAGATTTGTTTCAAAATGCTCTGTTGTAAGAGAG 322  
DB 181 TGTACCTTAAACTGTGATCAAGAAATTCCTATGGAGTTTGCCTCGGCGTTCAGAAAG 240  
QY 323 ATGATAAATAAACCCACTGATGCACCACTTTGTTGACAGGCTATCAGGGTTGCTACTAT 382  
DB 241 AAG-----AATGATCGATATGCACCACTGTTGCCAGGCACGAGGGTTGTAAGTAC 294  
QY 383 TTCGATCAAGATGGTATTTGTTGTAAGAGAGAGTCTGAAACCCAGACCACTGCT 442  
DB 295 TTCAGTGTATGGAACCTTTGTTGTAAGAGAGAGTCTGATCCTAGAAAATCCAAAGGCT 354  
QY 443 TAT 445  
DB 355 TGT 357

RESULT 8

US-10-157-622-1  
; Sequence 1, Application US/10157622  
; Publication No. US20030129720A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, Marilyn A.  
; Heath, Robyn L.  
; Clarke, Adrienne E.  
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scully, Scott, Murphy & Presser  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/157,622  
; FILING DATE: 29-May-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/454,295  
; FILING DATE: 01-SEP-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 9748  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; TELEX: 230 901 SANS UR

; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1104 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-157-622-1

Query Match 28.2%; Score 149; DB 15; Length 1104;  
Best Local Similarity 67.2%; Pred. No. 2.9e-32;  
Matches 244; Conservative 0; Mismatches 110; Indels 9; Gaps 2;  
QY 86 AAGCGTTGTACTAGAGAAATGTGTC---ATTTAGCTATGGCATATGCCACCGTTCAGAA 142  
DB 1 AAGCGTTGTACTTAACTGTGATCCAGAAATTCCTATGGAGTTTGCCTCGGCGTTCAGAA 60  
QY 143 GGAAGTCCCCAAAACCTATATGCACCAATTTGTTGCTCAGGCTATTAAGGGTTGCAACTAT 202  
DB 61 GAAAGAGAGATGATCGGATATGCACCACTGTTGCCAGGCACGAGGGTTGTAAGTAC 120  
QY 203 TACAGTGTCTAAGAGAGATTTGTTGTAAGAGAGATCTGACCTAGAAAACCCAAAGAT 262  
DB 121 TACAGTGTATGGAACCTTTGTTGTAAGAGAGATCTGATCCTAGAAAATCCAAAGGCT 180  
QY 263 TGTACCTTCGAATGTATACACAGATTTGTTTCAAAATGCTCTGTTGTAAGAGAGAAAG 322  
DB 181 TGTACCTTAAACTGTGATCAAGAAATTCCTATGGAGTTTGCCTCGGCGTTCAGAAAGAAAG 240  
QY 323 ATGATAAATAAACCCACTGATGCACCACTTTGTTGACAGGCTATCAGGGCTATCAGGGTTGCTACTAT 382  
DB 241 AAG-----AATGATCGATATGCACCACTGTTGCCAGGCACGAGGGTTGTAAGTAC 294  
QY 383 TTCGATCAAGATGGTATTTGTTGTAAGAGAGAGTCTGAAACCCAGACCACTGCT 442  
DB 295 TTCAGTGTATGGAACCTTTGTTGTAAGAGAGAGTCTGATCCTAGAAAATCCAAAGGCT 354  
QY 443 TAT 445  
DB 355 TGT 357

RESULT 9

US-10-072-809A-56  
; Sequence 56, Application US/10072809A  
; Publication No. US20030217382A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, Marilyn, A., Lay, Fung T., Heath, Robyn, L.  
; TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and us  
; TITLE OF INVENTION: therefor  
; FILE REFERENCE: 18-01  
; CURRENT APPLICATION NUMBER: US/10/072,809A  
; CURRENT FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: USSN 60/267,271  
; PRIOR FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 56  
; LENGTH: 1104  
; TYPE: DNA  
; ORGANISM: Nicotiana glauca  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1104)  
US-10-072-809A-56  
Query Match 28.2%; Score 149; DB 15; Length 1104;  
Best Local Similarity 67.2%; Pred. No. 2.9e-32;  
Matches 244; Conservative 0; Mismatches 110; Indels 9; Gaps 2;  
QY 86 AAGCGTTGTACTAGAGAAATGTGTC---ATTTAGCTATGGCATATGCCACCGTTCAGAA 142  
DB 1 AAGCGTTGTACTTAACTGTGATCCAGAAATTCCTATGGAGTTTGCCTCGGCGTTCAGAA 60

```
Db 1 AAGGCTTGTTACCTTAACCTGTGATCCAAAGAAATTGCCCTATGGAGTTTGCCCGCGCTTCAGAA 60
QY 143 GGAAGTCCCCAAAACCTATATGACCAAAATGTGTCTCAGGCTATAAGGGTTGCAACTAT 202
Db 61 GAAAGAGAGATGATCGGATATGACCACTGTGTGGCAGGCACGAGGGTTGTAGTAC 120
QY 203 TACAGTGTAAAGAGATTTGTTGTGAAGGAGAACTCTGACCTAGAAACCCAAAAGAT 262
Db 121 TTCAGTGTATGGAACCTTTTGTGTGAAGGAGAGTCTGATCTAGAAAATCCAAAGGCT 180
QY 263 TGTACCTTGAATGTGATACACAGATGCTTATCAAAATGTCTCTGTTTCAGAGGAAG 322
Db 181 TGTACCTTAAACTGTGATCCAAAGAAATGGCCTATGGAGTTTGGCCGCGTTCAGAGAAAG 240
QY 323 ATGATAAATTAACCCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 382
Db 241 AAG-----AATGATCGGATATGACCACTGTGTGGCAGGCACGAGGGTTGTAGTAC 294
QY 383 TTCGATCAAGATGGTGTATTTGTGTGAAGGAGAGTCTCTGAACCCCAAGACCACTGCT 442
Db 295 TTCAGTGTATGGAACCTTTTGTGTGAAGGAGAGTCTGATCTAGAAAATCCAAAGGCT 354
QY 443 TAT 445
Db 355 TGT 357

RESULT 10
US-10-027-632-281276
; Sequence 281276, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 281276
; LENGTH: 680
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-281276
Query Match 6.7%; Score 35.6; DB 13; Length 680;
Best Local Similarity 53.6%; Pred. No. 12;
Matches 74; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 246 CTAGAAACCCAAAGATTGACCTTGAATGTGATACACAGATTGCTTATTCAAAATGTC 305
Db 27 CTTACAGACCTTAAGAGATTGCTTCTTGGAAAATAAATACAGGTTTATAGGTATTTAGATGAC 86
QY 306 CTCGTTTCAGAGGAAGATGATTAATTAACCCACTGATGCACCACTTGTTCACGGGCT 365
Db 87 CTAATTAAGATATACAGTTATGAGAGGCTTGGTAAGCGGCTTGTGTGATGATGCTT 146
QY 366 ATCAGGGTTGCTACTATT 383
Db 147 TTAAGTTTACACCTTTT 164

RESULT 12
US-10-191-807-3/c
; Sequence 3, Application US/10191807
; Publication No. US20030068691A1
; GENERAL INFORMATION:
; APPLICANT: HU, Song et al.
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CL001275-PROV
; CURRENT APPLICATION NUMBER: US/10/191,807
; CURRENT FILING DATE: 2002-07-10
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 108359
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
```

```
Db 147 TTAAGTTTACACCTTTT 164

RESULT 11
US-10-027-632-281276
; Sequence 281276, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 281276
; LENGTH: 680
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-281276
Query Match 6.7%; Score 35.6; DB 15; Length 680;
Best Local Similarity 53.6%; Pred. No. 12;
Matches 74; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 246 CTAGAAACCCAAAGATTGACCTTGAATGTGATACACAGATTGCTTATTCAAAATGTC 305
Db 27 CTTACAGACCTTAAGAGATTGCTTCTTGGAAAATAAATACAGGTTTATAGGTATTTAGATGAC 86
QY 306 CTCGTTTCAGAGGAAGATGATTAATTAACCCACTGATGCACCACTTGTTCACGGGCT 365
Db 87 CTAATTAAGATATACAGTTATGAGAGGCTTGGTAAGCGGCTTGTGTGATGATGCTT 146
QY 366 ATCAGGGTTGCTACTATT 383
Db 147 TTAAGTTTACACCTTTT 164

RESULT 12
US-10-191-807-3/c
; Sequence 3, Application US/10191807
; Publication No. US20030068691A1
; GENERAL INFORMATION:
; APPLICANT: HU, Song et al.
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CL001275-PROV
; CURRENT APPLICATION NUMBER: US/10/191,807
; CURRENT FILING DATE: 2002-07-10
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 108359
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
```

```
; NAME/KEY: misc feature
; LOCATION: (1)...(108359)
; OTHER INFORMATION: n = A,T,C or G
US-10-191-807-3

Query Match      6.7%; Score 35.6; DB 14; Length 108359;
Best Local Similarity 45.9%; Pred. No. 1.6e+02;
Matches 122; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

QY 235 AGAATCTGACCCCTAGAACCCAAAGAGTTGTACCTTGGAAATGTGATACACAGATGCTTA 294
Db 850 ACAATATAACATAAATATATATTAATGATATATACCAACATATACACATAGCATG 791
QY 295 TTCAAAATGTCCTGGTTTCAGAAAGAGATGAATAATTAACCCACTGGATGCACCACTTG 354
Db 790 TGAAGGTATAAAAGATTACGTGAGTATGATAAAATACCAATCAGGATAATGGTTAAAGG 731
QY 355 TTCACGGGCTATCAGGGTTGCTACTATTTTCGATCAAGATGGTGATTTGCTGCTGAAGG 414
Db 730 TTTCACTGTGTTAAATGGTATTTAAATTTCTTAATCTAGGTGATTTTATTTCTTAATCT 671
QY 415 AGAGAGTCCTGAACCCCAAGACCACTGCTTATTTCTAATCAATCATATGTTGTTATCTATC 474
Db 670 AATGAACATATATACATGAATGTTTCATATATATATTTCTTCTTATAACTTCTATATGTT 611
QY 475 AAAAAAAATATCATGTCATGATATA 500
Db 610 CAAATTTATTCCTAATACTTTTTTA 585

RESULT 13
US-10-221-714A-285/c
; Sequence 285, Application US/10221714A
; Publication No. US20040048254A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; TITLE OF INVENTION: tumor suppressor genes and oncogenes
; FILE REFERENCE: 5013.1005
; CURRENT APPLICATION NUMBER: US/10/221,714A
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: PCT/EP01/02955
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: DE 10013847.0
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 540
; SEQ ID NO 285
; LENGTH: 6134
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-714A-285

Query Match      6.7%; Score 35.2; DB 16; Length 6134;
Best Local Similarity 44.9%; Pred. No. 49;
Matches 133; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

QY 220 TTTGATTTGTGAAGGAGAACTGACCCCTAGAAACCCCAAGAGTTGTACCTTCCGAATGTA 279
Db 4083 TTGCAATTTTACAAATAAATAAACAATTAATAATATAAATTTTAAACCTAA 4024
QY 280 TACACAGATTGCTTTATTCAAAATGTCTCTCGTCAGAGGAAGATGATATAATTAACCCAC 339
```

```
Db 4023 ATATAAAATTACATCTACATATATAACTCAATACAAAAAACAATATAATAAT 3964
QY 340 TGGATGCACCACTTGTTCACGGGCTATCAGGGTTGCTACTATTTTCGATCAAGATGGTA 399
Db 3963 AAAAATATATTAATAAAAAAACAATAAATACTATTATCTTTAAATATATCAAAATTTTC 3904
QY 400 TTTTGTCTGTGAAGGAGAGAGTCTGACCCAGACCACTGCTTATTTCTTAATCAATCAT 459
Db 3903 CTTTCTCTATTTTTCATTAACCTTCTAACACAAAAACCACTACTATTATATATAAAAAA 3844
QY 460 ATGTTGTTTCTATCAAAAAAATAATATGATGATGATATATATGCTGGTTACTGTAA 515
Db 3843 AATAATTTATTTTAACTAAACAATACTACTACTATAATCCCAAAACTTTTA 3788

RESULT 14
US-10-311-455-611
; Sequence 611, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 611
; LENGTH: 6650
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-611

Query Match      6.7%; Score 35.2; DB 15; Length 6650;
Best Local Similarity 68.1%; Pred. No. 51;
Matches 49; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 442 TTATTTCTTAATCAATCATATGTTGTTATCTATCAAAAAAATAATGTATCATGATATAT 501
Db 248 TTATTTTCTTTTATGTGATGTTTATGTATGTAATAAAAAATGTGTATGTTTGTATAG 307
QY 502 GCTGTTACTGT 513
Db 308 TTTGGATGGT 319

RESULT 15
US-10-311-455-352/c
; Sequence 352, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
```

;  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 2424  
; SEQ ID NO 352  
; LENGTH: 10433  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-311-455-352

Query Match 6.7%; Score 35.2; DB 15; Length 10433;  
Best Local Similarity 54.7%; Pred. No. 64;  
Matches 70; Conservative 0; Mismatches 58; Indels 0; Gaps 0;  
QY 130 CCCACGTTGAGAGGAGTCCCAAAAACCTATATGCACCAATTGTTGCTCAGGCTATAA 189  
Db 7458 CCAACCTCCATCTTAAAACTCTTAAAACTTTCAACCCACCAACATTTCTCTAAATATATAA 7399  
QY 190 GGGTTGCAACTATTACAGTCTAAAGGAGATTGATTGTGAAGGAGAACTCGACCCCTAG 249  
Db 7398 AAAAAACATCTCTCTCACTCTAAAAATACCTTTTTTTTAAATAAAAATCTCACTCTAT 7339  
QY 250 AAACCCAA 257  
Db 7338 TACCCAAA 7331

Search completed: January 22, 2005, 07:25:42  
Job time : 444 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2005, 04:53:35 ; Search time 83 Seconds  
(without alignments)  
4530.209 Million cell updates/sec

Title: US-10-725-829-1

Perfect score: 529  
Sequence: 1 cataatggctgttcacaag.....ctgtaatggactttattcg 529

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 35539441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	178.8	33.8	1360	3	US-08-454-295-2
2	178.8	33.8	1360	3	US-09-431-500A-2
3	178.8	33.8	1360	4	US-09-431-498-2
4	178.8	33.8	1360	4	US-09-431-499-2
5	149	28.2	1104	3	US-08-454-295-1
6	149	28.2	1104	3	US-09-431-500A-1
7	149	28.2	1104	3	US-09-431-498-1
8	149	28.2	1104	4	US-09-431-499-1
9	38.6	7.3	832	4	US-09-621-976-2813
10	35.8	6.8	1141	4	US-09-806-708B-22
11	34.8	6.6	6866	4	US-10-204-708-19
12	34.4	6.5	193303	4	US-09-497-855A-37
13	34.4	6.5	193303	4	US-09-497-855A-44
14	33.8	6.4	1141	4	US-09-806-708B-22
15	33.8	6.4	5473	4	US-09-620-312D-260
16	33.6	6.4	4043	4	US-09-710-279-3354
17	33.4	6.3	648	4	US-09-861-451A-25
18	33.2	6.3	2945	3	US-09-058-489-85
19	33.2	6.3	9439	3	US-09-058-489-89
20	33	6.2	777	4	US-09-540-236-1792
21	33	6.2	63563	4	US-09-596-002-33
22	32.8	6.2	12685	4	US-09-479-467A-3
23	32.6	6.2	558	4	US-09-134-000C-2786
24	32.6	6.2	1181	4	US-09-270-767-15127
25	32.4	6.1	832	4	US-09-621-976-2813
26	32	6.0	1314	4	US-09-248-796A-4870
27	31.8	6.0	1751	4	US-09-620-312D-847
					Sequence 2, Appl1
					Sequence 2, Appl1
					Sequence 2, Appl1
					Sequence 1, Appl1
					Sequence 1, Appl1
					Sequence 1, Appl1
					Sequence 1, Appl1
					Sequence 2813, Ap
					Sequence 22, Appl
					Sequence 19, Appl
					Sequence 37, Appl
					Sequence 44, Appl
					Sequence 22, Appl
					Sequence 260, App
					Sequence 3354, Ap
					Sequence 25, Appl
					Sequence 85, Appl
					Sequence 89, Appl
					Sequence 1792, Ap
					Sequence 33, Appl
					Sequence 3, Appl1
					Sequence 2786, Ap
					Sequence 15127, A
					Sequence 2813, Ap
					Sequence 4870, Ap
					Sequence 847, App

28 31.8 6.0 786431 4 US-09-751-389-3 Sequence 3, Appl1  
29 31.6 6.0 581 3 US-09-385-982-73 Sequence 73, Appl1  
30 31.6 6.0 9870 4 US-09-245-928A-15 Sequence 15, Appl1  
31 31.4 5.9 1338 3 US-09-134-001C-2111 Sequence 2111, Ap  
32 31.4 5.9 13737 3 US-09-538-414-10 Sequence 10, Appl1  
33 31.4 5.9 13737 4 US-10-074-279-10 Sequence 10, Appl1  
34 31.4 5.9 640681 4 US-09-790-288-1 Sequence 1, Appl1  
35 31.2 5.9 5262 4 US-09-248-796A-5051 Sequence 5051, Ap  
36 31 5.9 280 4 US-09-270-767-29958 Sequence 29958, A  
37 31 5.9 787 4 US-08-956-171B-584 Sequence 584, App  
38 31 5.9 787 4 US-08-781-986A-584 Sequence 13891, A  
39 31 5.9 1123 4 US-09-270-767-13891 Sequence 32, Appl1  
40 31 5.9 118057 4 US-09-497-855A-32 Sequence 25, Appl1  
41 30.8 5.8 585 3 US-08-671-548C-25 Sequence 10, Appl1  
42 30.8 5.8 639 3 US-09-196-293-10 Sequence 10, Appl1  
43 30.8 5.8 639 3 US-08-209-603E-10 Sequence 33, Appl1  
44 30.8 5.8 639 3 US-08-235-836C-33 Sequence 10, Appl1  
45 30.8 5.8 639 4 US-09-711-546-10 Sequence 10, Appl1

#### ALIGNMENTS

RESULT 1  
US-08-454-295-2  
; Sequence 2, Application US/08454295  
; Patent No. 6031087  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, Marilyn A.  
; APPLICANT: Atkinson, Angela H.  
; APPLICANT: Heath, Robyn L.  
; APPLICANT: Clarke, Adrienne E.  
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scully, Scott, Murphy & Presser  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/454,295  
; FILING DATE: 01-SEP-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 9748  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1360 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 97..1200  
US-08-454-295-2

Query Match 33.8%; Score 178.8; DB 3; Length 1360;  
Best Local Similarity 67.3%; Pred. No. 1.9e-46;  
Matches 303; Conservative 0; Mismatches 132; Indels 15; Gaps 3;



US-09-431-498-2

Query Match 33.8%; Score 178.8; DB 4; Length 1360;  
Best Local Similarity 67.3%; Pred. No. 1.9e-46;  
Matches 303; Conservative 0; Mismatches 132; Indels 15; Gaps 3;

QY 5 ATGGCTGTTTCACAAAGTTAGCTTCTTGGCTGCTACTTGTCTTGGATGGATGTTTCTA 64  
DB 10 ATGGCTGTTTCACAGAGTTAGTTCTTGGCTGCTCTCTCTTATTGGAAATGCTCTGCT 69  
QY 65 CT-----TGCAGAACATGTTGATGCCAAGGCTTGTACTAGAGAAATGTTGTCATT---TT 115  
DB 70 GTAAGCAATGTGGAACATGCAGATGCCAAGGCTTGTACCTTAACTGTGATCCAAAT 129  
QY 116 AGCTATGGCATATGCCCATGTTTGCAGAGGAAGTCCCAAAACCTATATGCACCAATGT 175  
DB 130 GCCTATGGAGTTTGCCTGCTTGCAGAGGAAGTCCCAAAACCTATATGCACCAATGT 189  
QY 176 TGCTCAGGCTATAGGGTTCACACTATTACAGTCTTAAGGAGATTTGATTGTGAAGGA 235  
DB 190 TGGCAGGCGACGAAGGTTGTAAGTACTTCAGTGATGATGGAACCTTTTGTGTGAAGGA 249  
QY 236 GAATCTGACCTAGAACCCCAAGATTTGACCTTGAATGTGATACACAGATTTGCTTAT 295  
DB 250 GAGTCTGATCTAGAAATCCAAAGGCTTGTACCTTAACTGTGATCCAAATTTGCTTAT 309  
QY 296 TCAAAATGTCCTGTTTGCAGAGGAAGATGATAATTAACCCCACTGGATGCCACCTTGT 355  
DB 310 GGAGTTTGCCTGCTTGCAGAGGAAGATGATAATTAACCCCACTGGATGCCACCTTGT 363  
QY 356 TGCAAGGCTATCAGGTTGCTACTATTTCGATCAAGATGGTGTATTTGTCTGTGAAGGA 415  
DB 364 TGGCAGGCGACGAAGGTTGTAAGTACTTCAGTGATGATGGAACCTTTTGTGTGAAGGA 423  
QY 416 GAGAGTCTTGAACCCCAAGACCTGCTTAT 445  
DB 424 GAGTCTGATCCTAGAAATCCAAAGGCTTGT 453

RESULT 4

US-09-431-499-2

/ Sequence 2, Application US/09431499  
/ Patent No. 6451573  
/ GENERAL INFORMATION:  
/ APPLICANT: Anderson, Marilyn A.  
/ APPLICANT: Atkinson, Angela H.  
/ APPLICANT: Heath, Robyn L.  
/ APPLICANT: Clarke, Adrienne E.  
/ TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF  
/ NUMBER OF SEQUENCES: 14  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: Scully, Scott, Murphy & Presser  
/ STREET: 400 Garden City Plaza  
/ CITY: Garden City  
/ STATE: New York  
/ COUNTRY: United States of America  
/ ZIP: 11530  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: PatentIn Release #1.0, Version #1.30  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/09/431,499  
/ FILING DATE:  
/ CLASSIFICATION:  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/454,295  
/ FILING DATE: 01-SEP-1995  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: DiGiglio, Frank S.  
/ REGISTRATION NUMBER: 31,346

REFERENCE/DOCKET NUMBER: 9748

TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 742-4343

TELEFAX: (516) 742-4366

TELEX: 230 901 SANS UR

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1360 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 97..1200

US-09-431-499-2

Query Match 33.8%; Score 178.8; DB 4; Length 1360;

Best Local Similarity 67.3%; Pred. No. 1.9e-46;

Matches 303; Conservative 0; Mismatches 132; Indels 15; Gaps 3;

QY 5 ATGGCTGTTTCACAAAGTTAGCTTCTTGGCTGCTACTTGTCTTGGATGGATGTTTCTA 64  
DB 10 ATGGCTGTTTCACAGAGTTAGTTCTTGGCTGCTCTCTCTTATTGGAAATGCTCTGCT 69  
QY 65 CT-----TGCAGAACATGTTGATGCCAAGGCTTGTACTAGAGAAATGTTGTCATT---TT 115  
DB 70 GTAAGCAATGTGGAACATGCAGATGCCAAGGCTTGTACCTTAACTGTGATCCAAAT 129  
QY 116 AGCTATGGCATATGCCCATGTTTGCAGAGGAAGTCCCAAAACCTATATGCACCAATGT 175  
DB 130 GCCTATGGAGTTTGCCTGCTTGCAGAGGAAGTCCCAAAACCTATATGCACCAATGT 189  
QY 176 TGCTCAGGCTATAGGGTTCACACTATTACAGTCTTAAAGGAGATTTGATTGTGAAGGA 235  
DB 190 TGGCAGGCGACGAAGGTTGTAAGTACTTCAGTGATGATGGAACCTTTTGTGTGAAGGA 249  
QY 236 GAATCTGACCTAGAACCCCAAGATTTGACCTTGAATGTGATACACAGATTTGCTTAT 295  
DB 250 GAGTCTGATCCTAGAAATCCAAAGGCTTGTACCTTAACTGTGATCCAAATTTGCTTAT 309  
QY 296 TCAAAATGTCCTGTTTGCAGAGGAAGATGATAATTAACCCCACTGGATGCCACCTTGT 355  
DB 310 GGAGTTTGCCTGCTTGCAGAGGAAGATGATAATTAACCCCACTGGATGCCACCTTGT 363  
QY 356 TGCAAGGCTATCAGGTTGCTACTATTTCGATCAAGATGGTGTATTTGTCTGTGAAGGA 415  
DB 364 TGGCAGGCGACGAAGGTTGTAAGTACTTCAGTGATGATGGAACCTTTTGTGTGAAGGA 423  
QY 416 GAGAGTCTTGAACCCCAAGACCTGCTTAT 445  
DB 424 GAGTCTGATCCTAGAAATCCAAAGGCTTGT 453

RESULT 5

US-08-454-295-1

/ Sequence 1, Application US/08454295

/ Patent No. 6031087

/ GENERAL INFORMATION:

/ APPLICANT: Anderson, Marilyn A.

/ APPLICANT: Atkinson, Angela H.

/ APPLICANT: Heath, Robyn L.

/ APPLICANT: Clarke, Adrienne E.

/ TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC

/ NUMBER OF SEQUENCES: 14

/ CORRESPONDENCE ADDRESS:

/ ADDRESSEE: Scully, Scott, Murphy &amp; Presser

/ STREET: 400 Garden City Plaza

/ CITY: Garden City

/ STATE: New York

/ COUNTRY: United States of America

/ ZIP: 11530

/ COMPUTER READABLE FORM:

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,295
; FILING DATE: 01-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1104 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-454-295-1

Query Match      28.2%; Score 149; DB 3; Length 1104;
Best Local Similarity 67.2%; Pred. No. 4.5e-37;
Matches 244; Conservative 0; Mismatches 110; Indels 9; Gaps 2;

QY      86 AAGGCTTGACTAGAGAAATGTGTC---ATTTAGCTATGGCATATGCCACGTTTCAGAA 142
Db      1 AAGGCTTGACTTAAACTGTGATCCAAGAAATTCGCTATGGAGTTTGGCCGCGTTTCAGAA 60

QY      143 GGAAGTCCCAAAACCTATATGCACCAATTTGTTGCTCAGGCTATTAAGGGTTTCAACTAT 202
Db      61 GAAAAGAAGAAATGATCGGATATGCACCAACTGTTTGGCAGGACGAAAGGGTTTGAAGTAC 120

QY      203 TACAGTGTCTAAAGGAGATTTGATTGCTGAAGGAGAAATCTGACCTAGAAACCCAAAGAT 262
Db      121 TTCAAGTATGATGGAACCTTTGTTGTTGAAGGAGAGTCTGATCTAGAAATCCAAAGGCT 180

QY      263 TGTACCTTCGAATGTGATACACAGATTGCTTATTTCAAAATGTCCTCGTTTCAGAAAGAAAG 322
Db      181 TGTACCTTAAACTGTGATCCAAGAAATTCGCTATGGAGTTTGGCCGCGTTTCAGAAAGAAAG 240

QY      323 ATGATAATTAACCCACTGGATGCACCACTTGTGTGACGGGCTATCAGGGTTGCTACTAT 382
Db      241 AAG-----AATGATCGGATATGCACCAACTGTTTGGCAGGACGAAAGGGTTTGAAGTAC 294

QY      383 TTCGATCAAGATGGTGTGATTGTTGCTGTGAAGGAGAGAGTCTCTGAACCCAGACCACTGCT 442
Db      295 TTCAGTGTATGGAACCTTTGTTGTTGAAGGAGAGTCTGATCTAGAAATCCAAAGGCT 354

443 TAT 445
355 TGT 357

RESULT 6
US-09-431-500A-1
; Sequence 1, Application US/09431500A
; Patent No. 6261821
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: IBM PC compatible
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/431,498
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: US/09/431,500A
; CURRENT FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 08/454,295
; ATTORNEY/AGENT INFORMATION:
```

```
; PRIOR FILING DATE: 1995-09-01
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1104
; TYPE: DNA
; ORGANISM: Nicotiana alata
; US-09-431-500A-1

Query Match      28.2%; Score 149; DB 3; Length 1104;
Best Local Similarity 67.2%; Pred. No. 4.5e-37;
Matches 244; Conservative 0; Mismatches 110; Indels 9; Gaps 2;

QY      86 AAGGCTTGACTAGAGAAATGTGTC---ATTTAGCTATGGCATATGCCACGTTTCAGAA 142
Db      1 AAGGCTTGACTTAAACTGTGATCCAAGAAATTCGCTATGGAGTTTGGCCGCGTTTCAGAA 60

QY      143 GGAAGTCCCAAAACCTATATGCACCAATTTGTTGCTCAGGCTATTAAGGGTTTCAACTAT 202
Db      61 GAAAAGAAGAAATGATCGGATATGCACCAACTGTTTGGCAGGACGAAAGGGTTTGAAGTAC 120

QY      203 TACAGTGTCTAAAGGAGATTTGATTGCTGAAGGAGAAATCTGACCTAGAAACCCAAAGAT 262
Db      121 TTCAAGTATGATGGAACCTTTGTTGTTGAAGGAGAGTCTGATCTAGAAATCCAAAGGCT 180

QY      263 TGTACCTTCGAATGTGATACACAGATTGCTTATTTCAAAATGTCCTCGTTTCAGAAAGAAAG 322
Db      181 TGTACCTTAAACTGTGATCCAAGAAATTCGCTATGGAGTTTGGCCGCGTTTCAGAAAGAAAG 240

QY      323 ATGATAATTAACCCACTGGATGCACCACTTGTGTGACGGGCTATCAGGGTTGCTACTAT 382
Db      241 AAG-----AATGATCGGATATGCACCAACTGTTTGGCAGGACGAAAGGGTTTGAAGTAC 294

QY      383 TTCGATCAAGATGGTGTGATTGTTGCTGTGAAGGAGAGAGTCTCTGAACCCAGACCACTGCT 442
Db      295 TTCAGTGTATGGAACCTTTGTTGTTGAAGGAGAGTCTGATCTAGAAATCCAAAGGCT 354

443 TAT 445
355 TGT 357

RESULT 7
US-09-431-498-1
; Sequence 1, Application US/09431498
; Patent No. 6440727
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: IBM PC compatible
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/431,498
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: 08/454,295
; CURRENT FILING DATE:
; ATTORNEY/AGENT INFORMATION:
```



```
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9748
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1104 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-431-498-1

Query Match      28.2%; Score 149; DB 4; Length 1104;
Best Local Similarity 67.2%; Pred. No. 4.5e-37;
Matches 244; Conservative 0; Mismatches 110; Indels 9; Gaps 2;

QY 86 AAGGCTTGTTACTAGAGAATGTGTC---ATTTTAGCTATGGCATATGCCACGTTTCAGAA 142
DB 1 AAGGCTTGTTACTTAACTGTGATCCAGAAATTCCTATGGAGTTTGCCTGTTTCAGAA 60
QY 143 GGAAGTCCCCAAAACCTATATGCACCAATTTGTTGCTCAGGCTATTAAGGTTTCAACTAT 202
DB 61 GAAAGAAGAATGATCGGATATGCACCACTGTTGGCAGGCACGAAGGGTTGTAAGTAC 120
QY 203 TACAGTGTCTGAATGTGATACACAGATTTGCTTATCAAAATGTCTCTGTTTCAGAAAGAAAG 322
DB 121 TTCACTTAACTGTGATCCAGAAATTCCTATGGAGTTTGCCTGTTTCAGAAAGAAAG 180
QY 263 TGTACCTTCGAATGTGATACACAGATTTGCTTATCAAAATGTCTCTGTTTCAGAAAGAAAG 322
DB 181 TGTACCTTAACTGTGATCCAGAAATTCCTATGGAGTTTGCCTGTTTCAGAAAGAAAG 240
QY 323 ATGATAATTAACCCACTGATGCACCACTGTTGTCAGGGCTATCAGGGTTGCTACTAT 382
DB 241 AAG-----AATGATCGGATATGCACCACTGTTGGCAGGCACGAAGGGTTGTAAGTAC 294
QY 383 TTCGATCAAGATGGTGTATTTGTTCTGTGAAGGAGAGAGTCTCTGAAACCCAGACCACTGCT 442
DB 295 TTCAGTGTATGATGGAACCTTTTGTGTAAGGAGAGTCTGATCCTAGAAATCCAAAGGCT 354
QY 443 TAT 445
DB 355 TGT 357

RESULT 8
US-09-431-499-1
Sequence 1, Application US/09431499
Patent No. 6451573
GENERAL INFORMATION:
APPLICANT: Anderson, Marilyn A.
APPLICANT: Atkinson, Angela H.
APPLICANT: Heath, Robyn L.
APPLICANT: Clarke, Adrienne E.
TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF
TITLE OF INVENTION: AND GENETIC
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States of America
ZIP: 11530
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/431,499
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/454,295
FILING DATE: 01-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9748
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1104 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-431-499-1

Query Match      28.2%; Score 149; DB 4; Length 1104;
Best Local Similarity 67.2%; Pred. No. 4.5e-37;
Matches 244; Conservative 0; Mismatches 110; Indels 9; Gaps 2;

QY 86 AAGGCTTGTTACTAGAGAATGTGTC---ATTTTAGCTATGGCATATGCCACGTTTCAGAA 142
DB 1 AAGGCTTGTTACTTAACTGTGATCCAGAAATTCCTATGGAGTTTGCCTGTTTCAGAA 60
QY 143 GGAAGTCCCCAAAACCTATATGCACCAATTTGTTGCTCAGGCTATTAAGGTTTCAACTAT 202
DB 61 GAAAGAAGAATGATCGGATATGCACCACTGTTGGCAGGCACGAAGGGTTGTAAGTAC 120
QY 203 TACAGTGTCTGAATGTGATACACAGATTTGCTTATCAAAATGTCTCTGTTTCAGAAAGAAAG 322
DB 121 TTCACTTAACTGTGATCCAGAAATTCCTATGGAGTTTGCCTGTTTCAGAAAGAAAG 240
QY 263 TGTACCTTCGAATGTGATACACAGATTTGCTTATCAAAATGTCTCTGTTTCAGAAAGAAAG 322
DB 181 TGTACCTTAACTGTGATCCAGAAATTCCTATGGAGTTTGCCTGTTTCAGAAAGAAAG 240
QY 323 ATGATAATTAACCCACTGATGCACCACTGTTGTCAGGGCTATCAGGGTTGCTACTAT 382
DB 241 AAG-----AATGATCGGATATGCACCACTGTTGGCAGGCACGAAGGGTTGTAAGTAC 294
QY 383 TTCGATCAAGATGGTGTATTTGTTCTGTGAAGGAGAGTCTCTGAAACCCAGACCACTGCT 442
DB 295 TTCAGTGTATGATGGAACCTTTTGTGTAAGGAGAGTCTGATCCTAGAAATCCAAAGGCT 354
QY 443 TAT 445
DB 355 TGT 357

RESULT 9
US-09-621-976-2813/c
Sequence 2813, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 2813
LENGTH: 832
```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
US-09-621-976-2813

Query Match
Best Local Similarity 14.1%; Pred. No. 0.028;
Matches 38; Conservative 116; Mismatches 115; Indels 0; Gaps 0;

QY 13 TCACAAAGTAGCTCTCTGCTTGCCTTACCTTCTGCTGATGATGTTCTACTTGGAA 72
Db 380 TAAATAATTATATTTTGTGTYTWWKTYWYWTYTRMMWKKKRRYYWKKSTYACAS 321

QY 73 ACATGTTGATGCCAAGCTTGATAGAGAATGTTGTTTGTGCTATGTCATGTCATGCCCC 132
Db 320 RYKRYTGWYWWYWKMMSTRWTCYWCCKCMYRGRCAYWTWARGMWSYANGKKSMR 261

QY 133 ACCTTCAGAGGAAGTCCCCAAAACCTATATACCAATTTGTTGCTCAGGCTATAAGG 192
Db 260 SAMSMCTRMYYKKGSTVWTKTCATWCYWKYKRWMSKTCWSGSRGGYMTSYTSTRS 201

QY 193 TTCCAACTATTACAGTCTAAAGAGATTGATTGTTGAAGGAGAATCTGACCCCTAGAA 252
Db 200 YSMYASWMTWCMWGRWWSYTYWAWGKKWRYATTWRRAMWMAAWTMMWYMWNAW 141

QY 253 CCAAAAGATTGTACCTTCGAATGTGATA 281
Db 140 CMSRGAAMYRRITMMWGHYWRKKSIR 112

RESULT 10
US-09-806-708B-22/c
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAEL promoters
US-09-806-708B-22

Query Match
Best Local Similarity 11.5%; Pred. No. 0.26;
Matches 44; Conservative 159; Mismatches 180; Indels 2; Gaps 1;

QY 96 CTAGAGAATGTGCTATTTTAGCTATGGCATATGCCCATGTCAGGCTATACAGTCTAAAG 215
Db 581 YTRNTYCKSYAHSYWWSNNMYRRYSARNSSMARWTTTRNNWWSGBVRWRWAGTW 522

QY 156 AACCTATATGCACCAATTTGTTGCTAGGCTATAGGGTTGCTCAACTATTACAGTCTAAAG 215
Db 521 WRHNNNTDTRYWYWKRWARETTTTYDSMCNAKSMWRGNWNRWAKMWWAANNADAGMD 462

QY 216 GAGATTGTTGTTGAAGGAGAATCTGACCTAGAACCCAAAAGATTGTACCTCGAAT 275
Db 461 HTWYMGNNNTMMRWAKMMNMAWCCRAYCCNNNNRACVWHKHKWRWTKWTKWKAACN 402

QY 276 GTGATACACAGATTGCTTATTCAAATGTCTCTGTTTCTGAGGAAGGAGATGATAATTAAC 335

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
US-09-621-976-2813

Query Match
Best Local Similarity 14.1%; Pred. No. 0.028;
Matches 38; Conservative 116; Mismatches 115; Indels 0; Gaps 0;

QY 13 TCACAAAGTAGCTCTCTGCTTGCCTTACCTTCTGCTGATGATGTTCTACTTGGAA 72
Db 380 TAAATAATTATATTTTGTGTYTWWKTYWYWTYTRMMWKKKRRYYWKKSTYACAS 321

QY 73 ACATGTTGATGCCAAGCTTGATAGAGAATGTTGTTTGTGCTATGTCATGTCATGCCCC 132
Db 320 RYKRYTGWYWWYWKMMSTRWTCYWCCKCMYRGRCAYWTWARGMWSYANGKKSMR 261

QY 133 ACCTTCAGAGGAAGTCCCCAAAACCTATATACCAATTTGTTGCTCAGGCTATAAGG 192
Db 260 SAMSMCTRMYYKKGSTVWTKTCATWCYWKYKRWMSKTCWSGSRGGYMTSYTSTRS 201

QY 193 TTCCAACTATTACAGTCTAAAGAGATTGATTGTTGAAGGAGAATCTGACCCCTAGAA 252
Db 200 YSMYASWMTWCMWGRWWSYTYWAWGKKWRYATTWRRAMWMAAWTMMWYMWNAW 141

QY 253 CCAAAAGATTGTACCTTCGAATGTGATA 281
Db 140 CMSRGAAMYRRITMMWGHYWRKKSIR 112

RESULT 11
US-10-204-708-19
; Sequence 19, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 19
; LENGTH: 6866
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-19

Query Match
Best Local Similarity 52.0%; Pred. No. 1.4;
Matches 78; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 380 TATTTCGATCAAGATGTCATTTTGTCTGTGAGGAGAGAGTCTCTGAACCCAGACCACT 439
Db 1648 TAGTTGGCGGATGAGGTGAGATTGTTTAAATAAAAAAAAAAAAAAAAAAAAAAT 1707

QY 440 GCTTATTCTCAATCATATGTTCTATCATCAAAAAAAAAAATATGTCATGATGATAT 499
Db 1708 TTTTATTAGATTGAAGTTTATTATTAGGATATAAATAATGTTAGTTATATTT 1767

QY 500 ATGCTGGTTACTGTAAATGTGGACTTTATTG 529
Db 1768 GTTATAATTTATAAAGTTTATTTTATTAG 1797

RESULT 12
US-09-497-855A-37
; Sequence 37, Application US/09497855A
; Patent No. 6605432
; GENERAL INFORMATION:
; APPLICANT: Huang, Tim
; TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR DETECTING DNA METHYLATION
; FILE REFERENCE: UMO1523
```

```
; CURRENT APPLICATION NUMBER: US/09/497,855A
; CURRENT FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/120,592
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: 60/118,760
; PRIOR FILING DATE: 1999-02-05
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 37
; LENGTH: 193303
; TYPE: DNA
; ORGANISM: Homo sapiens;
US-09-497-855A-37

Query Match      6.5%; Score 34.4; DB 4; Length 193303;
Best Local Similarity 52.9%; Pred. No. 11;
Matches 74; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 218 GATTTCATTGTCAGGAGATCTGACCCCTAGAAACCCAAAGATGTTACCTTCGAATGT 173233
DB 173174 GACTTCAGAGCTATTGGAGAAATCAGAGGGTAAATTCACAGAAATCTCTAAATTTAAATGT 173233

QY 278 GATACACAGATGCTTATTCAAAATGTCCTCGTTTCAGAGGAAGATGATAATTAAACCC 337
DB 173234 GACTAAAGTTTATATTCTTCAGTTCTGCTCATGAAGTTGAGAGAAATAGATAAATGCG 173293

QY 338 ACTGGATGCACCACTTGTTG 357
DB 173294 ATTCACGGAGCATTTGTGG 173313

RESULT 13
US-09-497-855A-44
; Sequence 44, Application US/09497855A
; Patent No. 6605432
; GENERAL INFORMATION:
; APPLICANT: Huang, Tim
; TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR DETECTING DNA METHYLATION
; FILE REFERENCE: UMO1523
; CURRENT APPLICATION NUMBER: US/09/497,855A
; CURRENT FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/120,592
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: 60/118,760
; PRIOR FILING DATE: 1999-02-05
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 44
; LENGTH: 193303
; TYPE: DNA
; ORGANISM: Homo sapiens;
US-09-497-855A-44

Query Match      6.5%; Score 34.4; DB 4; Length 193303;
Best Local Similarity 52.9%; Pred. No. 11;
Matches 74; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 218 GATTTCATTGTCAGGAGATCTGACCCCTAGAAACCCAAAGATGTTACCTTCGAATGT 173233
DB 173174 GACTTCAGAGCTATTGGAGAAATCAGAGGGTAAATTCACAGAAATCTCTAAATTTAAATGT 173233

QY 278 GATACACAGATGCTTATTCAAAATGTCCTCGTTTCAGAGGAAGATGATAATTAAACCC 337
DB 173234 GACTAAAGTTTATATTCTTCAGTTCTGCTCATGAAGTTGAGAGAAATAGATAAATGCG 173293

QY 338 ACTGGATGCACCACTTGTTG 357
DB 173294 ATTCACGGAGCATTTGTGG 173313

RESULT 14
US-09-806-708B-22
; Sequence 22, Application US/09806708B
```

```
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: Consensus sequence of A.t., L.a., and B.n. FAE1 promoters
US-09-806-708B-22

Query Match      6.4%; Score 33.8; DB 4; Length 1141;
Best Local Similarity 11.1%; Pred. No. 1.1;
Matches 40; Conservative 136; Mismatches 185; Indels 0; Gaps 0;

QY 131 CCAGTTCAGAGGAAGTCCCAAAACCTATATGACCAATTTGTTGCTCAGGCTATAAG 190
DB 345 VYNNNNNNNTYKKARHBARDWVHSAWKWHANAHAHSRKMTBYKRKTVMVNNNGT 404

QY 191 GGTTCGAACATTTACAGTCTAAAGGAGATTTGATTTGTGAAGGAGAGATCTGACCCCTAGA 250
DB 405 TWMKRWAWYWKMDMDWBGTYNNNNNGRTYTGTTGKTKMMWYTKKANNCKRWANDHKT 464

QY 251 AACCCAAAAGATTTGACCTTCGAATGTGATACACAGATTCCTTATTCAAAATGTCCTCGT 310
DB 465 CTHNNTTWMKMTYNNCYKSWTNGSKHRBAAAVVTYWMWRRYAHANNNDYWKKA 524

QY 311 TCAGAGGAAGATGATAATTAACCCACTGGATGACCACTTTGTCACGGGCTATCAG 370
DB 525 CTWYKYBVCWKWNNYAAYTKSSWNTSRYVYRWKTNNSWRWSRSDTRSMGRANNYARABH 584

QY 371 GGTTCGTACTATTTCGATCAAGATGCTGATTTGTCTGTGAAGGAGAGATCTCTGAACCC 430
DB 585 YGYKMNTRWBWSHTWBHBRAGAAHYMMBMWYBAKCHWKAWYKAKKYAGAGGSGNNNNN 644

QY 431 AAGACCACCTGTTATTCTTAATCAATCATATGTTGTTATCTATCAAAAAAATATGTAT 490
DB 645 NNNNNNNNATCARDYYAASRWYANAKWYVYKBAANNAAYYTHANNWGCWNNATDT 704

QY 491 G 491
DB 705 R 705

RESULT 15
US-09-620-312D-260
; Sequence 260, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenchua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunging
; APPLICANT: Wang, Dunrui
```

; APPLICANT: Wang, Zhiwei  
; APPLICANT: John Tillinghast  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 784CIP2B  
; CURRENT APPLICATION NUMBER: US/09/620,312D  
; CURRENT FILING DATE: 2000-07-19  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: pt FL\_genes Version 1.0  
; SEQ ID NO 260  
; LENGTH: 5473  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(4482)  
US-09-620-312D-260

Query Match 6.4%; Score 33.8; DB 4; Length 5473;  
Best Local Similarity 48.7%; Pred. No. 2.5;  
Matches 92; Conservative 0; Mismatches 97; Indels 0; Gaps 0;  
Qy 125 ATATGCCACGTTTCAGNAGGAAGTCCCAAAACCTATATGCACCAATTGTTGCTCAGGC 184  
Db 4988 ATATGATCACATGTAGAAGTAGCTTCAAGAAATTTCTGTAGTCATAAATGTTTAATAATA 5047  
Qy 185 TATAAGGGTTGCAACTATTATACAGTGCTAAAGGAGATTGATTTGTGAAGGAGAACTCTGAC 244  
Db 5048 TATGATGTAAATATATATATATGAGCCTATATGATGATGATGATGATGATGATGATGATGAT 5107  
Qy 245 CCTAGAAACCCAAAGATTCTACCTTCGAATGTGATACACAGATTGCTTATTCAAAATGT 304  
Db 5108 CTTAGAAAATGTAGACATGTTAACTGGGAAATATAAATATAGAGTGGCACTTCAAGACAA 5167  
Qy 305 CCTCGTTCA 313  
Db 5168 GCTGACTCA 5176

Search completed: January 22, 2005, 07:18:15  
Job time : 88 secs



Db 1 MAVHKEVNFAYLLIVLGNFL---YVDAAKATRECQNLGFGICPRSEGSPLNPICINCCS 57  
 Qy 60 GYKGCNYYSAKGDLLICEGSDPRNPCKDTFECDTQIAYSKCPRSEGMIIKPTGCTTCCT 119  
 Db 58 GYKGCNYYNSFGKFCICEGSDPKRPNACTFNCDPNIAYSRCPRSQGKSLIVPTGCTTCCT 117  
 Qy 120 GYQGCYFPDQDGFVCEGESPEPKTAY 147  
 Db 118 GYKGCYTFGDKGFVCEGESDEPKANMY 145  
 RESULT 4  
 T07597  
 proteinase inhibitor II - potato  
 C/Species: Solanum tuberosum (potato)  
 C/Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 09-Jul-2004  
 C/Accession: T07597  
 R/Lee, J.S.  
 submitted to the EMBL Data Library, June 1992  
 A/Reference number: S24965  
 A/Accession: T07597  
 A/Status: preliminary; translated from GB/EMBL/DDBJ  
 A/Molecule type: DNA  
 A/Residues: 1-158 <LBE>  
 A/Cross-references: UNIPROT:Q00782; EMBL:Z12753; NID:g21553; PID:CAA78277.1; PID:g21554  
 A/Experimental source: cv. Russet Burbank  
 C/Genetics:  
 A/Introns: 17/1  
 C/Superfamily: potato proteinase inhibitor PTI  
 C/Keywords: serine proteinase inhibitor  
 Query Match 69.4%; Score 593.5; DB 2; Length 158;  
 Best Local Similarity 71.9%; Pred. No. 1.3e-41;  
 Matches 105; Conservative 12; Mismatches 26; Indels 3; Gaps 2;  
 Qy 1 MAVHK-VSFLACLVLVGMFLA--KHVDAAKATRECQHFSGYICPRSEGSPPQPICTNC 57  
 Db 1 MAIHKEVSFLAYLLVLGMLLFVSAMEHVDAKACTLECNLGYGICPRSEGSPPENICTNC 60  
 Qy 58 CSYGKGCNYYSAKGDLLICEGSDPRNPCKDTFECDTQIAYSKCPRSEGMIIKPTGCTTC 117  
 Db 61 CAGYKGCNYYSAANGTFICEGSHPRNPKACPRNCDPHIAYSRCPRSGGKTLIVPTGCTTC 120  
 Qy 118 CTGYGCGCYFPDQDGFVCEGESPEPK 143  
 Db 121 CTGYTDCYTFGDKGFVCEGESIEPK 146  
 RESULT 5  
 S24973  
 proteinase inhibitor II - potato  
 C/Species: Solanum tuberosum (potato)  
 C/Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
 C/Accession: S24973  
 R/Choi, Y.; Kim, J.W.; Lee, J.S.  
 submitted to the EMBL Data Library, July 1992  
 A/Description: Characterization of a potato proteinase inhibitor II gene that is expressed  
 A/Reference number: S24973  
 A/Accession: S24973  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-147 <CHO>  
 A/Cross-references: UNIPROT:Q41489; EMBL:Z13992; NID:g21555; PID:CAA78383.1; PID:g21556  
 C/Genetics:  
 A/Introns: 18/1  
 C/Superfamily: potato proteinase inhibitor PTI  
 Query Match 68.5%; Score 586; DB 2; Length 147;  
 Best Local Similarity 72.2%; Pred. No. 5.1e-41;  
 Matches 104; Conservative 12; Mismatches 24; Indels 4; Gaps 2;  
 Qy 1 MAVHK-VSFLACLVLVGMFLA--KHVDAAKATRECQHFSGYICPRSEGSPPQPICTNC 59

```
Db 1 MAVHKEVSVAYLLVLVGMFL---YVDALCTKEGNGLFGICPRSEGSPTNFCINCCS 57
QY 60 GYKGCNYSAGDLICGESDPRNPDKCTFECDTQIAYSKCPRSEGMIIKPTGCTTC 119
Db 58 GYKGCNYSAGDLICGESDPRNPDKCTFECDTQIAYSKCPRSEGMIIKPTGCTTC 117
QY 120 GYKGCNYSAGDLICGESDPRNPDKCTFECDTQIAYSKCPRSEGMIIKPTGCTTC 117
Db 118 GYKGCNYSAGDLICGESDPRNPDKCTFECDTQIAYSKCPRSEGMIIKPTGCTTC 117

RESULT 6
XRPOCI
proteinase inhibitor PCI-I precursor - potato
N;Alternate names: potato chymotrypsin inhibitor I; protease inhibitor II
C;Species: Solanum tuberosum (potato)
C;Date: 06-Jul-1992 #sequence_revision 12-Apr-1996 #text_change 09-Jul-2004
C;Accession: A26584; A01319; A23591
R;Thornburg, R.W.; An, G.; Cleveland, T.E.; Johnson, R.; Ryan, C.A.
Proc. Natl. Acad. Sci. U.S.A. 84, 744-748, 1987
A;Title: Wound-inducible expression of a potato inhibitor II-chloramphenicol acetyltransferase
A;Reference number: A26584
A;Accession: A26584
A;Molecule type: DNA
A;Residues: 1-153 <THO>
A;Cross-references: UNIPROT:P01080; GB:M15186; NID:g169488; PIDN:AAA33815.1; PID:g169489
R;Hass, G.M.; Hermodson, M.A.; Ryan, C.A.; Gentry, L.
Biochemistry 21, 752-756, 1982
A;Title: Primary structures of two low molecular weight proteinase inhibitors from potato
A;Reference number: A90465; MUID:82182863; PMID:7074039
A;Accession: A01319
A;Molecule type: protein
A;Residues: 55-106 <HAS>
A;Note: Leu-92 is probably the site of interaction with chymotrypsin
R;Keil, M.; Sanchez-Serrano, J.; Schell, J.; Willmitzer, L.
Nucleic Acids Res. 14, 5641-5650, 1986
A;Title: Primary structure of a proteinase inhibitor II gene from potato (Solanum tuberosum)
A;Reference number: A23591; MUID:86286579; PMID:3016659
A;Accession: A23591
A;Molecule type: DNA
A;Residues: 1-26, 'EH', 28-33, 'TL', 36-54, 'R', 56-153 <KEI>
A;Cross-references: GB:X04118; NID:g21521; FIDN:CAA27730.1; PID:g21522
C;Genetics:
A;Gene: IIK
A;Intons: 18/1
A;Superfamily: potato proteinase inhibitor PTI
C;Keywords: serine proteinase inhibitor
F;1-25/Domain: signal sequence #status predicted <SIG>
F;55-106/Product: proteinase inhibitor PCI-I #status experimental <MAT>

Query Match 68.5%; Score 585.5; DB 1; Length 153;
Best Local Similarity 70.0%; Pred. No. 5.7e-41;
Matches 105; Conservative 14; Mismatches 28; Indels 3; Gaps 3;

QY 1 MAVHK-VSLFACL-LVLGMFL--AKHVDKACTREG-HFSYGICPRSEGSPOKPICTNC 57
Db 1 MDVHKEVNFVAYLLVLVGLLVLSVMDVDKACIRECGNLGFGICPRSEGSPTNFCIN 60
QY 58 CSYGKGCNYSAGDLICGESDPRNPDKCTFECDTQIAYSKCPRSEGMIIKPTGCTTC 117
Db 61 CAGYKGCNYSANGAFICEQSDPKPKKACPLNCDPHIAYSKCPRSEGSPTNFCIN 120
QY 118 CTGYGCGYFFDODGDFVCEGESPEPKTTAY 147
Db 121 CTGYKGCYFFGKNGKFCVCEGESDEPKANNY 150

RESULT 7
S43105
proteinase inhibitor II - potato
C;Species: Solanum tuberosum (potato)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
```

```
C;Accession: S43105
R;Murray, C.; Christeller, J.T.
submitted to the EMBL Data Library, March 1994
A;Reference number: S43105
A;Accession: S43105
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-154 <MUR>
A;Cross-references: UNIPROT:Q43652; EMBL:X78275; NID:g467609; PIDN:CAA55082.1; PID:g46761
C;Genetics:
A;Intons: 18/1
C;Superfamily: potato proteinase inhibitor PTI

Query Match 68.0%; Score 581; DB 2; Length 154;
Best Local Similarity 67.5%; Pred. No. 1.3e-40;
Matches 102; Conservative 16; Mismatches 29; Indels 4; Gaps 2;

QY 1 MAVHK-VSLFACL-LVLGMFL--AKHVDKACTREG-HFSYGICPRSEGSPOKPICTNC 56
Db 1 MDVHKEVNFVAYLLVLVGLLVSVVHVDKAKICTKEGNLGFGICPRSEGSPTNFCIN 60
QY 57 CSYGKGCNYSAGDLICGESDPRNPDKCTFECDTQIAYSKCPRSEGMIIKPTGCTTC 116
Db 61 CSYGKGCNYSVFGRFICEGSDLNKPKACPLNCDTNIAYSRCPHSEGSPTNFCIN 120
QY 117 CTGYGCGYFFDODGDFVCEGESPEPKTTAY 147
Db 121 CTGYKGCYFFGKNGKFCVCEGESDEPKANNY 151

RESULT 8
JQ2153
proteinase inhibitor II precursor - Persian tobacco
C;Species: Nicotiana glauca (Persian tobacco)
C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004
C;Accession: JQ2153; PQ0647; S65396
R;Atkinson, A.H.; Heath, R.L.; Simpson, R.J.; Clarke, A.E.; Anderson, M.A.
Plant Cell 5, 203-213, 1993
A;Title: Proteinase inhibitors in Nicotiana glauca stigmas are derived from a precursor p
A;Reference number: JQ2153; MUID:93200805; PMID:8453302
A;Accession: JQ2153
A;Molecule type: mRNA
A;Residues: 1-397 <ATK>
A;Cross-references: UNIPROT:Q40378; GB:U08219; NID:g473590; PIDN:AAA17739.1; PID:g473591
A;Accession: PQ0647
A;Molecule type: protein
A;Residues: 54-66; 112-124; 170-182; 228-240; 286-298; 344-357 <AT2>
A;Experimental source: stigma, style
R;Heath, R.L.; Barton, P.A.; Simpson, R.J.; Reid, G.E.; Lim, G.; Anderson, M.A.
Eur. J. Biochem. 230, 250-257, 1995
A;Title: Characterization of the protease processing sites in a multidomain proteinase p
A;Reference number: S65396; MUID:95324532; PMID:7601108
A;Accession: S65396
A;Molecule type: protein
A;Residues: 54-227; 286-343 <HEA>
A;Comment: This protein contains six similar domains, each with a potential active site.
C;Superfamily: potato proteinase inhibitor PTI
C;Keywords: serine proteinase inhibitor
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-397/Product: proteinase inhibitor II #status predicted <MAT>
F;35-36, 93-94/Region: chymotrypsin-specific sites
F;153-154, 211-212, 269-270, 327-328/Region: trypsin-specific sites

Query Match 52.6%; Score 450; DB 2; Length 397;
Best Local Similarity 57.5%; Pred. No. 1.3e-29;
Matches 84; Conservative 14; Mismatches 42; Indels 6; Gaps 4;

QY 1 MAVHKVSVFLACL-LVLGMFL--AKHVDKACTREG-HFSYGICPRSEGSPOKPICTNC 57
Db 1 MAVHRSVFLALLLLFCMSLLVSVNHADAKACTLNCDPIAYGVCPSEKKNDRICNTC 60
QY 58 CSYGKGCNYSAGDLICGESDPRNPDKCTFECDTQIAYSKCPRSEGMIIKPTGCTTC 117
Db 117 CSYGKGCNYSAGDLICGESDPRNPDKCTFECDTQIAYSKCPRSEGMIIKPTGCTTC 117
```





QY 53 ICTNCCSGYKGNYSYAKGLDICEGSDPRNPKDCTFECDTQIAYSKCPR 102  
Db 2 ICTNCCAGYKGCNYSGANGAFICEGSDPKNPVCPNCDNTIAYSKCLR 51

RESULT 13

JQ2269  
trypsin inhibitor-1 - common tobacco  
N/Alternate names: proteinase inhibitor  
C/Species: Nicotiana tabacum (common tobacco)  
C/Date: 19-May-1994 #sequence\_revision 19-May-1994 #text\_change 09-Jul-2004  
C/Accession: JQ2269  
R/Pearce, G.; Johnson, S.; Ryan, C.A.  
Plant Physiol. 102, 639-644, 1993  
A/Title: Purification and characterization from tobacco (Nicotiana tabacum) leaves of si  
A/Reference number: JQ2269; PMID:94151442; PMID:8108514  
A/Accession: JQ2269  
A/Molecule type: protein  
A/Residues: 1-53 <PEA>  
A/Cross-references: UNIPROT:Q7M1P5  
A/Experimental source: leaf  
A/Superfamily: potato proteinase inhibitor PTI  
C/Keywords: serine proteinase inhibitor

Query Match 23.3%; Score 199; DB 2; Length 53;  
Best Local Similarity 64.7%; Pred. No. 8e-10;  
Matches 33; Conservative 4; Mismatches 14; Indels 0; Gaps 0;

QY 53 ICTNCCSGYKGNYSYAKGLDICEGSDPRNPKDCTFECDTQIAYSKCPRS 103  
Db 3 ICTNCCAGYKGCYFSDDTGTFVCEGSDPRNPKACPNCDDPRIAYGICPLS 53

RESULT 14

TIB01  
proteinase inhibitor - eggplant  
C/Species: Solanum melongena (eggplant, aubergine)  
C/Date: 30-Nov-1980 #sequence\_revision 30-Nov-1980 #text\_change 09-Jul-2004  
C/Accession: A01317  
R/Richardson, M.  
FEBS Lett. 104, 322-326, 1979  
A/Title: The complete amino acid sequence and the trypsin reactive (inhibitory) site of  
A/Reference number: A01317; PMID:80041163; PMID:477995  
A/Accession: A01317  
A/Molecule type: protein  
A/Residues: 1-52 <RIC>  
A/Cross-references: UNIPROT:P01078  
A/Note: 2-Leu and 6-Cys were also found  
C/Superfamily: potato proteinase inhibitor PTI  
C/Keywords: pyroglutamic acid  
F/1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F/38/Inhibitory site: Arg (trypsin) #status experimental

Query Match 20.5%; Score 175; DB 1; Length 52;  
Best Local Similarity 58.8%; Pred. No. 7.1e-08;  
Matches 30; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 53 ICTNCCSGYKGNYSYAKGLDICEGSDPRNPKDCTFECDTQIAYSKCPRS 103  
Db 2 ICTNCCAGYKGCYFSDDTGTFVCEGSDPRNPKACPNCDDPRIAYGICPLS 52

RESULT 15

XKPO2A  
proteinase inhibitor IIa - potato (fragment)  
C/Species: Solanum tuberosum (potato)  
C/Date: 02-Apr-1982 #sequence\_revision 02-Apr-1982 #text\_change 09-Jul-2004  
C/Accession: A01320  
R/Iwasaki, T.; Kiyohara, T.; Yoshikawa, M.  
J. Biochem. 79, 181-191, 1976  
A/Title: Amino acid sequence of an active fragment of potato proteinase inhibitor IIa.  
A/Reference number: A01320; PMID:76190064; PMID:1270410  
A/Accession: A01320

A/Molecule type: protein  
A/Residues: 1-45 <IWA>  
A/Cross-references: UNIPROT:P01081  
A/Note: this active fragment inhibits trypsin strongly and chymotrypsin temporarily; the  
C/Superfamily: potato proteinase inhibitor PTI  
C/Keywords: serine proteinase inhibitor  
F/10-24, 14-35, 20-43/Disulfide bonds: #status experimental  
F/32/Inhibitory site: Lys (trypsin) #status experimental

Query Match 14.9%; Score 127; DB 1; Length 45;  
Best Local Similarity 57.1%; Pred. NO. 0.00052;  
Matches 24; Conservative 4; Mismatches 10; Indels 4; Gaps 1;

QY 45 SEGSPQKPICTNCCSGYKGCNYYS----AKGDLICEGSDPR 82  
Db 1 SEGSPENRICTNCCAGYKGCNYNDNTIAYSKVCEGFDPK 42

Search completed: January 24, 2005, 16:52:48  
Job time : 38 secs

THIS PAGE IS BLANK

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2005, 16:43:10 ; Search time 192 Seconds

(without alignments)  
443.518 Million cell updates/sec

Title: US-10-725-829-2

Perfect score: 855

Sequence: 1 MAVHKVSFLACLLVLGMFL.....QDGFVCEGESPEPKTTAYF 148

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 02:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	855	100.0	148	Q8W2K6	Q8W2K6 solanum ame
2	614.5	71.9	197	IP21_TOBAC	Q40561 nicotiana t
3	612	71.6	152	Q6JKN9	Q6JKN9 solanum nig
4	612	71.6	152	AAR37362	AAR37362 solanum n
5	602	70.4	201	IP23_LYCES	Q43502 lycopersico
6	596	69.7	148	IP21_LYCES	P05119 lycopersico
7	596	69.7	148	AAG12170	AAG12170 lycopersi
8	596	69.7	148	BAC76901	BAC76901 lycopersi
9	596	69.7	148	AAN05016	AAN05016 lycopersi
10	595	69.6	147	IP2T_SOLTU	Q41435 solanum tub
11	595	69.6	154	Q84156	Q84156 solanum phu
12	595	69.6	154	AAS01731	AAS01731 solanum p
13	593.5	69.4	158	IP2X_SOLTU	Q00782 solanum tub
14	589	68.9	154	IP25_SOLTU	Q41488 solanum tub
15	586	68.5	147	IP2Y_SOLTU	Q41489 solanum tub
16	585.5	68.5	153	IP2K_SOLTU	P01080 solanum tub
17	581	68.0	154	IP27_SOLTU	Q43652 solanum tub
18	555	64.9	126	Q82735	Q82735 solanum tub
19	487	57.0	109	Q8W2K1	Q8W2K1 solanum ame
20	450	52.6	397	Q40378	Q40378 nicotiana a
21	449	52.5	281	Q9S077	Q9S077 nicotiana a
22	440	51.5	455	Q84L39	Q84L39 nicotiana a
23	430	50.3	204	IP22_CAPAN	Q49146 capsicum an
24	425.5	49.8	506	Q9SDW8	Q9SDW8 nicotiana g
25	425	49.7	204	Q9SDL4	Q9SDL4 capsicum an
26	416	48.7	390	Q9SDW7	Q9SDW7 nicotiana g
27	343.5	40.2	410	Q6H238	Q6H238 nicotiana a
28	334	39.1	146	Q7M222	Q7M222 lycopersico
29	333	38.9	223	IP22_LYCES	Q43710 lycopersico
30	245	28.7	55	IP21_CAPAN	P56615 capsicum an
31	231.5	27.1	143	Q9M514	Q9M514 capsicum an

RESULT 1

ID	Q8W2K6	PRELIMINARY;	PRT;	148 AA.
AC	Q8W2K6;			
DT	01-MAR-2002 (Tremblrel. 20, Created)			
DT	01-MAR-2002 (Tremblrel. 20, Last sequence update)			
DT	01-JUN-2003 (Tremblrel. 24, Last annotation update)			
DE	Proteinase inhibitor IIA.			
GN	NamesPINza;			
OS	Solanum americanum.			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;			
OC	lamiids; Solanales; Solanaceae; Solanum.			
OX	NCBI_TaxID=109975;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21643209; PubMed=11785934;			
RA	Xu Z.-F., Qi W.-O., Ouyang X.-Z., Yeung E., Chye M.-L.;			
RT	"A proteinase inhibitor II of Solanum americanum is expressed in			
RT	phloem.";			
RL	Plant Mol. Biol. 47:727-738 (2001).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Chye M.-L., Xu Z.-F.;			
RL	Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF174381; AAL36458.1; -.			
DR	HSSP; P05119; IOYV.			
DR	GO; GO:0004867; P:serine-type endopeptidase inhibitor activity; IEA.			
DR	InterPro; IPR003465; Prot inh Pin2.			
DR	Pfam; PF02428; Prot_inhib_II; 2.			
SQ	SEQUENCE 148 AA, 15323 MW, 053536378E104E75 CRC64;			

Query Match 100.0%; Score 855; DB 2; Length 148;

Best Local Similarity 100.0%; Pred. No. 1.1e-69;

Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAVHKVSFLACLLVLGMFLAKHDAKCTRECGHFSYGCIPRSGSPKPICTNCCSG	60
Db	1	MAVHKVSFLACLLVLGMFLAKHDAKCTRECGHFSYGCIPRSGSPKPICTNCCSG	60
Qy	61	YKGCNYSAGDILCEGESDPRNPDKCTFECDTQIAYSKCPRSEGKMIKPTGCTCTG	120
Db	61	YKGCNYSAGDILCEGESDPRNPDKCTFECDTQIAYSKCPRSEGKMIKPTGCTCTG	120
Qy	121	YQCYFYFDQDGFVCEGESPEPKTTAYF	148
Db	121	YQCYFYFDQDGFVCEGESPEPKTTAYF	148

RESULT 2

IP21_TOBAC				
ID	IP21_TOBAC	STANDARD;	PRT;	197 AA.
AC	Q40561;			
DT	13-DEC-1998 (Rel. 37, Created)			

DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Proteinase inhibitor type II precursor.  
 OS Nicotiana tabacum (Common tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC lamids; Solanales; Solanaceae; Nicotiana.  
 OX NCBI\_TaxID=4097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Sameun NN; TISSUE=Leaf;  
 RX MEDLINE=95284369; PubMed=776901;  
 RA Balandin M.T., van der Does C., Albert J.M., Bol J.F.,  
 RA Linthorst H.J.M.;  
 RT "Structure and induction pattern of a novel proteinase inhibitor class  
 RT II gene of tobacco.";  
 RL Plant Mol. Biol. 27:1197-1204(1995).  
 CC -I- INDUCTION: Locally induced in leaves subjected to different types  
 CC of stress (TMV infection, wounding, UV irradiation).  
 CC -I- SIMILARITY: Belongs to the potato type II proteinase inhibitor  
 CC family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z29537; CAA82652.1; -;  
 DR PIR; S56662; S56662.  
 DR HSP; P05119; IOYV.  
 DR InterPro; IPR003465; Prot\_inh\_Pin2.  
 DR Pfam; PF02428; Prot\_inhib\_II; 3.  
 KW Repeat; Serine protease inhibitor; Signal.  
 FT SIGNAL 1 24 Potential.  
 FT CHAIN 25 197 Proteinase inhibitor type II.  
 FT REPEAT 24 80 1.  
 FT REPEAT 81 140 2.  
 FT REPEAT 141 196 3.  
 FT SITE 29 30 Reactive bond for trypsin (Potential).  
 FT DISULFID 27 115 By similarity.  
 FT DISULFID 31 111 By similarity.  
 FT DISULFID 39 121 By similarity.  
 FT DISULFID 51 88 By similarity.  
 FT DISULFID 54 72 By similarity.  
 FT DISULFID 55 84 By similarity.  
 FT DISULFID 61 97 By similarity.  
 FT DISULFID 114 132 By similarity.  
 SQ SEQUENCE 197 AA; F9CD5CB4267A9710 CRC64;  
 Query Match 71.9%; Score 614.5; DB 1; Length 197;  
 Best Local Similarity 75.7%; Pred. No. 8.2e-48;  
 Matches 109; Conservative 12; Mismatches 20; Indels 3; Gaps 2;  
 QY 1 MAVHKVSFLACLVLGWMFLAKHVDKACTREGHFSYGCIPRSEGSPOKPICTNCSG 60  
 DB 1 MAVHKVSFVAHLVLVLG-MFLULL--VDKAKTKCGNFAYGICPRSQGTDPDPICTTCCAG 57  
 QY 61 YKGCNYSYSAKGLICEGSDPRNPKDCTFECDTQIAYSKCPSESGKMIKPTGCTTCCG 120  
 DB 58 YKGCNYSYANGTICGSSDPKPNVCPQFCDDPIAYSKCPSEGETIINPTGCTTCCG 117  
 QY 121 YQGYTFDQDGFVCEGSPEPKT 144  
 DB 118 YKGCYTFQDGFVCEGSDEPKS 141  
 RESULT 3  
 Q6JKN9 PRELIMINARY; PRT; 152 AA.  
 ID Q6JKN9  
 AC Q6JKN9;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Proteinase inhibitor 2b precursor.  
 GN Name=PIN2b;  
 OS Solanum nigrum (Black nightshade).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC lamids; Solanales; Solanaceae; Solanum.  
 OX NCBI\_TaxID=4112;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=15078438;  
 RA Schmidt D.D., Kessler A., Kessler D., Schmidt S., Lim M., Gase K.,  
 RA Baldwin I.T.;  
 RT "Solanum nigrum: a model ecological expression system and its tools.";  
 RL Mol. Ecol. 13:981-995(2004).  
 DR EMBL; AY422686; AAR37362.1; -. Pin2.  
 DR InterPro; IPR003465; Prot\_inh\_Pin2.  
 DR Pfam; PF02428; Prot\_inhib\_II; 2.  
 KW Signal.  
 FT SIGNAL 1 31 Potential.  
 SQ SEQUENCE 152 AA; 16432 MW; D86DD64DCD46125C CRC64;  
 Query Match 71.6%; Score 612; DB 2; Length 152;  
 Best Local Similarity 74.7%; Pred. No. 1.1e-47;  
 Matches 109; Conservative 9; Mismatches 24; Indels 4; Gaps 2;  
 QY 1 MAVHK-VSFLACLVLGWMFLLA---KHVDKACTREGHFSYGCIPRSEGSPOKPICTN 56  
 DB 1 MAVHKEVSSLAIVLLVGLLFLFVSAIKHVDKAPCTREGCNLGYGICPRSEGSPEINICTN 60  
 QY 57 CCGYKGCNYSYSAKGLICEGSDPRNPKDCTFECDTQIAYSKCPSESGKMIKPTGCTT 116  
 DB 61 CCGYKGCNYSYANGTICGSSDPKPNVCPYCDGDIAYSKCPSESGOTIIVPGCTT 120  
 QY 117 CTGYQGYCYFDDQDGFVCEGSPEP 142  
 DB 121 CTGYKGCYVFSKEGFEVCEGSIEP 146  
 RESULT 4  
 AAR37362 PRELIMINARY; PRT; 152 AA.  
 ID AAR37362  
 AC AAR37362;  
 DT 24-MAY-2004 (TrEMBLrel. 27, Created)  
 DT 24-MAY-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 24-MAY-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Proteinase inhibitor 2b precursor.  
 GN PIN2B.  
 OS Solanum nigrum (Black nightshade).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC lamids; Solanales; Solanaceae; Solanum.  
 OX NCBI\_TaxID=4112;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=15078438;  
 RA Schmidt D.D., Kessler A., Kessler D., Schmidt S., Lim M., Gase K.,  
 RA Baldwin I.T.;  
 RT "Solanum nigrum: a model ecological expression system and its tools.";  
 RL Mol. Ecol. 13:981-995(2004).  
 DR EMBL; AY422686; AAR37362.1; -.  
 KW Signal.  
 FT SIGNAL 1 31 Potential.  
 SQ SEQUENCE 152 AA; 16432 MW; D86DD64DCD46125C CRC64;  
 Query Match 71.6%; Score 612; DB 2; Length 152;  
 Best Local Similarity 74.7%; Pred. No. 1.1e-47;  
 Matches 109; Conservative 9; Mismatches 24; Indels 4; Gaps 2;  
 QY 1 MAVHK-VSFLACLVLGWMFLLA---KHVDKACTREGHFSYGCIPRSEGSPOKPICTN 56  
 DB 1 MAVHKEVSSLAIVLLVGLLFLFVSAIKHVDKAPCTREGCNLGYGICPRSEGSPEINICTN 60  
 QY 57 CCGYKGCNYSYSAKGLICEGSDPRNPKDCTFECDTQIAYSKCPSESGKMIKPTGCTT 116  
 DB 61 CCGYKGCNYSYANGTICGSSDPKPNVCPYCDGDIAYSKCPSESGOTIIVPGCTT 120

Db 1 MAVHKEVSSLAYLLVLGLLFLVSAIKHVDKPTRECGNLGYGICPRSESGSPENPCTN 60  
 Qy 57 CCGYKGCNYYSAKGLICGSDPRNPXDCCTFECDTQIAYSKPRSEGMIIKPTGCTT 116  
 Db 61 CCGYKGCNYYSAKGLICGSDPRNPXDCCTFECDTQIAYSKPRSEGMIIKPTGCTT 120  
 Qy 117 CCGYKGCNYYSAKGLICGSDPRNPXDCCTFECDTQIAYSKPRSEGMIIKPTGCTT 142  
 Db 121 CCGYKGCNYYSAKGLICGSDPRNPXDCCTFECDTQIAYSKPRSEGMIIKPTGCTT 146

## RESULT 5

IP23\_LYCES STANDARD; PRT; 201 AA.  
 AC Q43502;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Proteinase inhibitor type II CEVI57 precursor.  
 GN Names=CEVI57;  
 OS Lycopersicon esculentum (Tomato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC Lamiales; Solanales; Solanaceae; Solanum.  
 OX NCBI\_TaxID=4081;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Rutgers; TISSUE=Leaf;  
 RC MEDLINE=96252900; PubMed=8672818;  
 RA Gadea J., Mayda E., Conejero V., Vera P.;  
 RT "Characterization of defense-related genes ectopically expressed in  
 RT viroid-infected tomato plants.";  
 RL Mol. Plant Microbe Interact. 9:409-415 (1996).  
 CC -I- INDUCTION: By viroid infection.  
 CC -I- SIMILARITY: Belongs to the potato type II proteinase inhibitor  
 CC family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X94946; CAA64416.1; -.  
 DR PIR; T07011; T07011.  
 DR HSSP; P05119; IOYV.  
 DR InterPro; IPR003465; Prot\_inh\_Pin2.  
 DR Pfam; PF02428; Prot\_inh\_II\_3.  
 KW Repeat; Serine protease inhibitor; Signal.

FT SIGNAL 1 23 Potential.  
 FT CHAIN 24 201 Proteinase inhibitor type II CEVI57.  
 FT REPEAT 27 83 1.  
 FT REPEAT 84 143 2.  
 FT REPEAT 144 199 3.  
 FT SITE 32 33 Reactive bond for trypsin (Probable).  
 FT SITE 147 148 Reactive bond for trypsin (Probable).  
 FT DISULFID 30 118 By similarity.  
 FT DISULFID 34 114 By similarity.  
 FT DISULFID 42 124 By similarity.  
 FT DISULFID 54 91 By similarity.  
 FT DISULFID 57 75 By similarity.  
 FT DISULFID 58 87 By similarity.  
 FT DISULFID 64 100 By similarity.  
 FT DISULFID 117 135 By similarity.  
 SQ SEQUENCE 201 AA; 21419 MW; A3FCAB93773D8590 CRC64;

Query Match 70.4%; Score 602; DB 1; Length 201;  
 Best Local Similarity 75.7%; Pred. No. 1.le-46;  
 Matches 109; Conservative 9; Mismatches 24; Indels 2; Gaps 2;

Qy 1 MAVHKEVSSLAYLLVLGLLFLVSAIKHVDKPTRECGNLGYGICPRSESGSPENPCTN 59

Db 1 MAVHKEVSSLAYLLVLGLLFLVSAIKHVDKPTRECGNLGYGICPRSESGSPENPCTN 59  
 Qy 60 GYKGCNYYSAKGLICGSDPRNPXDCCTFECDTQIAYSKPRSEGMIIKPTGCTT 119  
 Db 60 GYKGCNYYSAKGLICGSDPRNPXDCCTFECDTQIAYSKPRSEGMIIKPTGCTT 119  
 Qy 120 GYKGCNYYSAKGLICGSDPRNPXDCCTFECDTQIAYSKPRSEGMIIKPTGCTT 143  
 Db 120 GYKGCNYYSAKGLICGSDPRNPXDCCTFECDTQIAYSKPRSEGMIIKPTGCTT 146

## RESULT 6

IP21\_LYCES STANDARD; PRT; 148 AA.  
 AC P05119;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 01-OCT-2004 (Rel. 45, Last annotation update)  
 DE Wound-induced proteinase inhibitor II precursor.  
 OS Lycopersicon esculentum (Tomato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC Lamiales; Solanales; Solanaceae; Solanum.  
 OX NCBI\_TaxID=4081;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Leaf;  
 RC MEDLINE=85207658; PubMed=3838986;  
 RA Graham J.S., Pearce G., Merryweather J., Titani K., Ericsson L.H.,  
 RA Ryan C.A.;  
 RT "Wound-induced proteinase inhibitors from tomato leaves. II. The cDNA-  
 RT deduced primary structure of pre-inhibitor II.";  
 RL J. Biol. Chem. 260:6561-6564 (1985).  
 RN [2]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 31-141 IN COMPLEX WITH  
 RP SUBTILISIN.  
 RX PubMed=12684499; DOI=10.1074/jbc.M302020200;  
 RA Barrette-Ng I.H., Ng K.K.-S., Cherney M.M., Pearce G., Ryan C.A.,  
 RA James M.N.G.;  
 RT "Structural basis of inhibition revealed by a 1:2 complex of the two-  
 RT headed tomato inhibitor-II and subtilisin Carlsberg.";  
 RL J. Biol. Chem. 278:24062-24071 (2003).  
 CC -I- FUNCTION: Potent inhibitor of both trypsin and chymotrypsin.  
 CC -I- SUBCELLULAR LOCATION: Secreted.  
 CC -I- INDUCTION: Mechanical damage (i.e., insect chewing) to this plant  
 CC results in the systemic release of a factor from the wound site.  
 CC Within the leaves it induces the cytoplasmic synthesis of  
 CC proteinase inhibitors I and II.  
 CC -I- SIMILARITY: Belongs to the potato type II proteinase inhibitor  
 CC family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; K03291; AAA34201.1; -.  
 DR PIR; B24048; B24048.  
 DR PDB; IOYV; X-ray; I=26-148.  
 DR PDB; 1PJU; X-ray; A/B/C/D=26-148.  
 DR InterPro; IPR003465; Prot\_inh\_Pin2.  
 DR Pfam; PF02428; Prot\_inh\_II\_2.  
 KW 3D-structure; Repeat; Serine protease inhibitor; Signal.

FT SIGNAL 1 25  
 FT CHAIN 26 148 Wound-induced proteinase inhibitor II.  
 FT REPEAT 26 81 1 (trypsin-inhibitory).  
 FT REPEAT 83 141 2 (chymotrypsin-inhibitory).  
 FT SITE 30 31 Reactive bond for trypsin.  
 FT SITE 87 88 Reactive bond for chymotrypsin.



```

RESULT 11
Q84L56 PRELIMINARY; PRT; 154 AA.
ID Q84L56;
AC AC Q84L56;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE DE Proteinase inhibitor II protein (Proteinase inhibitor 2 protein).
GS Names=PIN-II2x; Synonyms=Bn-I12x;
OS Solanum phureja.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=172790;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Wound-induced leaf;
RA Bu Q., Yang S., Wan J.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
[2]
RN SEQUENCE FROM N.A.
RC TISSUE=Wound-induced leaf;
RA Bu Q., Wu L., Yang S., Wan J.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
DR ENBL; AY247794; AAC08244.1; -
DR DR ENBL; AY517498; AAS01731.1; -
DR DR HSP; P05119; IOYV.
DR GO; GO:0004867; F:serine-type endopeptidase inhibitor activity; IEA.
DR InterPro; IPR003465; Prot_inh_Pln2.
DR Pfam; PF02428; Prot_inhib_II_2.
DR DR SEQUENCE 154 AA; 16641 MW; F7019E715BF58553 CRC64;
SQ
Query Match 69.6%; Score 595; DB 2; Length 154;
Best Local Similarity 70.9%; Pred. No. 3.8e-46;
Matches 107; Conservative 15; Mismatches 25; Indels 4; Gaps 3;

Qy 1 MAVHK-VSLACL-LVLGMFLAA--KHVDKAKTRECQHFSYGICPRSEGSQKPICITN 56
Db 1 MAVHKEVNFVAYLLIVGLLLVLVSAMEHVDAKACTLECNLGFICPRSEGSPENRICTN 60
Qy 57 CCGSYGKGNYSNAKDILCEGESDPNPXDKCTPECDTQTATSKCPRSEGHWIIPKTGCTT 116
Db 61 CCAGYKGCNYNSANGAFICEGESDPPKNVPCPNCDNTIAYSKCPRSEGKSLIYPTGCTT 120
Qy 117 CCTGYGCVXYFDDGDGVFCGESEPEPKTTAY 147
Db 121 CCTGYKGCYFGNGKXFVCEGESDEPKANNY 151

RESULT 12
AAS01731 PRELIMINARY; PRT; 154 AA.
ID AAS01731
AC AAS01731;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE DE Proteinase inhibitor 2 protein.
GS PIN-II2X.
OS Solanum phureja.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=172790;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. IVP101; TISSUE=Wound-induced leaf;
RA Bu Q., Wu L., Yang S., Wan J.;
RL "Diploid potato proteinase inhibitor.";
RL Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
DR ENBL; AY517498; AAS01731.1; -
DR DR SEQUENCE 154 AA; 16641 MW; F7019E715BF58553 CRC64;
SQ

```





```
Db 1 MAVHKEVNFVAYLLIIVGLVLVLSAMEHVDAKACTLECGNLGFGICPRSEGSPEINICTN 60
Qy 57 CCSGYKGCNYSAGDLICEGSDPRNPKDCTFECTQIAYSKCPSESGKMIKPTGCTT 116
Db 61 CCAGYKGCNYSANGAFICEGSDPKKPKACPRNCDPHIAYSKCPSESGKSLIYPTGCTT 120
Qy 117 CCTGYGCGCYFFDQDGFVCEGESPEPKTAY 147
Db 121 CCTGYGCGCYFFGNGKGFVCEGESDEPKANNY 151
```

## RESULT 15

```
IP2Y_SOLTU STANDARD; PRT; 147 AA.
AC Q41489;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Proteinase inhibitor type II precursor.
OS Solanum tuberosum (potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamiales; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Russet Burbank;
RA Choi Y., Kim J.W., Lee J.S.;
RL Submitted (JUL-1992) to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: Belongs to the potato type II proteinase inhibitor
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z13992; CAA78383.1; -
DR FIR; S24973; S24973.
DR HSP; P05119; 10YV.
DR InterPro; IPR003465; Prot inh Pin2.
DR Pfam; PF02428; Prot inh II; 2.
KW Repeat; Serine protease inhibitor; Signal.
FT SIGNAL 1 25 Potential.
FT CHAIN 26 147 Proteinase inhibitor type II.
FT REPEAT 25 81 1.
FT REPEAT 82 141 2.
FT SITE 30 31 Reactive bond for trypsin (Probable).
FT SITE 87 88 Reactive bond for chymotrypsin
FT (Probable).
FT DISULFID 28 116 By similarity.
FT DISULFID 32 112 By similarity.
FT DISULFID 40 122 By similarity.
FT DISULFID 52 89 By similarity.
FT DISULFID 55 73 By similarity.
FT DISULFID 56 85 By similarity.
FT DISULFID 62 98 By similarity.
FT DISULFID 115 133 By similarity.
SQ SEQUENCE 147 AA; 15936 MW; 703456551B54F968 CRC64;
```

```
Query Match 58.5%; Score 586; DB 1; Length 147;
Best Local Similarity 72.2%; Pred. No. 2.4e-45;
Matches 104; Conservative 12; Mismatches 24; Indels 4; Gaps 2;
Qy 1 MAVHK-VSFLACLIVLGMWFLAKHVDKACTRECCHFSYVICPRSEGSPOKPICTNCCS 59
Db 1 MAVHKEVNFVAYLLIIVGLMFL----YDALGCTKECGNLGFGICPRSEGSPTNPICNCCS 57
Qy 60 GYKGCNYSAGDLICEGSDPRNPKDCTFECTQIAYSKCPSESGKMIKPTGCTTCT 119
```

```
Db 58 GYKGCNYSAGRFICEGESDPNPKACPFNCDTNIAYSRCPSESGKSLIYPTGCTTCT 117
Qy 120 GYQGCYFFDQDGFVCEGESPEPK 143
Db 118 GYKGCYFFGTNGKGFVCEGESDEPK 141
```

Search completed: January 24, 2005, 16:52:05  
Job time : 194 secs

**THIS PAGE IS BLANK**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2005, 16:44:21 ; Search time 40 Seconds  
(without alignments)  
245.377 Million cell updates/sec

Title: US-10-725-829-2  
Perfect score: 855  
Sequence: 1 MAVHKVSFLACLLVGLWFL.....QGDGFVCEGSEPEPKTTAYF 148

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	379	44.3	368	3	US-08-454-295-3
2	379	44.3	368	3	US-09-431-500A-3
3	379	44.3	368	4	US-09-431-498-3
4	379	44.3	368	4	US-09-431-499-3
5	224	26.2	58	3	US-08-454-295-5
6	224	26.2	58	3	US-09-431-500A-5
7	224	26.2	58	4	US-09-431-498-5
8	224	26.2	58	4	US-09-431-499-5
9	204	23.9	58	3	US-08-454-295-6
10	204	23.9	58	3	US-09-431-500A-6
11	204	23.9	58	4	US-09-431-498-6
12	204	23.9	58	4	US-09-431-499-6
13	203	23.7	58	3	US-08-454-295-7
14	203	23.7	58	3	US-08-454-295-8
15	203	23.7	58	3	US-09-431-500A-7
16	203	23.7	58	4	US-09-431-500A-8
17	203	23.7	58	4	US-09-431-498-7
18	203	23.7	58	4	US-09-431-498-8
19	203	23.7	58	4	US-09-431-499-7
20	203	23.7	58	4	US-09-431-499-8
21	199	23.3	58	3	US-08-454-295-9
22	199	23.3	58	3	US-09-431-500A-9
23	199	23.3	58	4	US-09-431-498-9
24	199	23.3	58	4	US-09-431-499-9
25	120	14.0	54	3	US-08-454-295-10
26	120	14.0	54	3	US-09-431-500A-10
27	120	14.0	54	4	US-09-431-498-10

28 120 14.0 54 4 US-09-431-499-10 Sequence 10, Appl  
29 95.5 11.2 578 3 US-08-981-392-13 Sequence 13, Appl  
30 95.5 11.2 578 4 US-09-908-322-13 Sequence 13, Appl  
31 94 11.0 148 3 US-08-882-907-15 Sequence 15, Appl  
32 94 11.0 148 4 US-10-032-658-15 Sequence 15, Appl  
33 92.5 10.8 2321 4 US-09-230-652-2 Sequence 2, Appl  
34 91.5 10.7 242 4 US-09-312-283C-393 Sequence 193, App  
35 91.5 10.7 420 4 US-09-907-794A-109 Sequence 109, App  
36 91.5 10.7 420 4 US-09-905-125A-109 Sequence 109, App  
37 91.5 10.7 420 4 US-09-902-775A-109 Sequence 109, App  
38 91.5 10.7 420 4 US-09-906-700-109 Sequence 290, App  
39 91.5 10.7 420 4 US-10-140-002-290 Sequence 109, App  
40 91.5 10.7 420 4 US-09-903-603A-109 Sequence 109, App  
41 91 10.6 1417 3 US-08-900-230-3 Sequence 3, Appl  
42 89.5 10.5 314 2 US-08-460-309-19 Sequence 19, Appl  
43 89.5 10.5 314 2 US-08-125-077-19 Sequence 19, Appl  
44 89.5 10.5 2213 1 US-08-727-034-3 Sequence 3, Appl  
45 89.5 10.5 2214 1 US-08-727-034-7 Sequence 7, Appl

## ALIGNMENTS

RESULT 1  
US-08-454-295-3  
; Sequence 3, Application US/08454295  
; Patent No. 6031087  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, Marilyn A.  
; APPLICANT: Atkinson, Angela H.  
; APPLICANT: Heath, Robyn L.  
; APPLICANT: Clarke, Adrienne E.  
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scully, Scott, Murphy & Presser  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/454,295  
; FILING DATE: 01-SEP-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 9748  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 368 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-454-295-3

Query Match 44.3%; Score 379; DB 3; Length 368;  
Best Local Similarity 58.1%; Pred. No. 2.3e-27;  
Matches 68; Conservative 9; Mismatches 36; Indels 4; Gaps 3;  
Qy 28 KACTRECG-HFSGICPRSEGSQKPCITNCSCGYKCNYSKAGDLICBEGSDPRPKD 86  
Db 1 KACTLNCDPRIAYGVCPREKKNDRICNCCAGTKGCKYFSDDGTFVCEGSDPRPKA 60

```

Qy      87 CTPECTQTAYSKCPRSEGMKIPTGCTTCTCTGGVCGCYFFDQGDGFVCEGESPEPK 143
Db      61 CTLNCDPRIAYGVCPRSEKK--NDRICTNCAGTKGCKYFSDDGTFVCEGES-DPR 114

RESULT 2
US-09-431-500A-3
; Sequence 3, Application US/09431500A
; Patent No. 6261821
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; TITLE OF INVENTION: SEQUENCES ENCODING SAME
; FILE REFERENCE: 9748B
; CURRENT APPLICATION NUMBER: US/09/431,500A
; CURRENT FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 08/454,295
; PRIOR FILING DATE: 1995-09-01
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Nicotiana alata
US-09-431-500A-3

Query Match          44.3%; Score 379; DB 3; Length 368;
Best Local Similarity 58.1%; Pred. No. 2.3e-27; Indels 4; Gaps 1
Matches 68; Conservative 9; Mismatches 36;

Qy      28 KACTRECG-HFSYGICPRSEGSPQKPICTNCCSGYGCNNYSAGKDLCIGESDPRNPKD 86
Db      1  KACTLNCDPRIAYGVCPRSEKKNDRICTNCCAGTKGCKYFSDDGTFVCEGESDPRNPKA 60

Qy      87 CTPECTQTAYSKCPRSEGMKIPTGCTTCTCTGGVCGCYFFDQGDGFVCEGESPEPK 143
Db      61 CTLNCDPRIAYGVCPRSEKK--NDRICTNCAGTKGCKYFSDDGTFVCEGES-DPR 114

RESULT 3
US-09-431-498-3
; Sequence 3, Application US/09431498
; Patent No. 6440727
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/431,498
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/454,295
; FILING DATE:

```

```
; MOLECULE TYPE: protein
US-09-431-499-3

Query Match      44.3%; Score 379; DB 4; Length 368;
Best Local Similarity 58.1%; Pred. No. 2.3e-27;
Matches 68; Conservative 9; Mismatches 36; Indels 4; Gaps 3;

Qy 28 KACTREG-HFSGICPRSGSQKPICTNCCSGYKGNYSKAGDLICEGSDPRNPKD 86
Db 1 KACTLNCDPRIAYGVCPRSEKKNDRICTNCCAGTKGCKYFSDGTFVCEGESDPRNPKA 60

Qy 87 CTPECDTQIAYSKPRSEGMIIKPTGCTTCTCTGYGCGYFDDQDFVCEGESPEPK 143
Db 61 CTLNCDPRIAYGVCPRSEKK--NDRICTNCCAGTKGCKYFSDGTFVCEGES-DPR 114

RESULT 5
US-08-454-295-5
; Sequence 5, Application US/08454295
; Patent No. 6031087
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,295
; FILING DATE: 01-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 58 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-454-295-5

Query Match      26.2%; Score 224; DB 3; Length 58;
Best Local Similarity 68.5%; Pred. No. 6.9e-14;
Matches 37; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

Qy 53 ICTNCCSGYKGNYSKAGDLICEGSDPRNPKDCTFECDTQIAYSKCPRSECK 106
Db 3 ICTNCCAGTKGCKYFSDGTFVCEGESDPRNPKACTLNCDPRIAYGVCPRSECK 56

RESULT 6
US-09-431-500A-5
; Sequence 5, Application US/09431500A
; Patent No. 6261821
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/431,498
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/454,295
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 58 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-431-500A-5

Query Match      26.2%; Score 224; DB 3; Length 58;
Best Local Similarity 68.5%; Pred. No. 6.9e-14;
Matches 37; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

Qy 53 ICTNCCSGYKGNYSKAGDLICEGSDPRNPKDCTFECDTQIAYSKCPRSECK 106
Db 3 ICTNCCAGTKGCKYFSDGTFVCEGESDPRNPKACTLNCDPRIAYGVCPRSECK 56

RESULT 7
US-09-431-498-5
; Sequence 5, Application US/09431498
; Patent No. 6440727
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/431,498
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/454,295
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 58 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-431-500A-5

Query Match      26.2%; Score 224; DB 3; Length 58;
Best Local Similarity 68.5%; Pred. No. 6.9e-14;
Matches 37; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

Qy 53 ICTNCCSGYKGNYSKAGDLICEGSDPRNPKDCTFECDTQIAYSKCPRSECK 106
Db 3 ICTNCCAGTKGCKYFSDGTFVCEGESDPRNPKACTLNCDPRIAYGVCPRSECK 56
```

```
; MOLECULE TYPE: protein
US-09-431-498-5
Query Match      26.2%; Score 224; DB 4; Length 58;
Best Local Similarity 68.5%; Pred. No. 6.9e-14;
Matches 37; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 53 ICTNCCSGYKGNYSAGDLICEGSDPRNPDKCTFECDTQIAYSKCPRSEK 106
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 3 ICTNCCAGTKGCKYFSDGTFVCEGESDPRNPACTLNCDPRIAYGVCPRSEK 56

RESULT 8
US-09-431-499-5
; Sequence 5, Application US/09431499
; Patent No. 6451573
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF
; TITLE OF INVENTION: AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/431,499
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,295
; FILING DATE: 01-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELEPHONE: (516) 742-4366
; TELEFAX: (516) 742-4343
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 58 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-431-499-5
Query Match      26.2%; Score 224; DB 4; Length 58;
Best Local Similarity 68.5%; Pred. No. 6.9e-14;
Matches 37; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 53 ICTNCCSGYKGNYSAGDLICEGSDPRNPDKCTFECDTQIAYSKCPRSEK 106
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 3 ICTNCCAGTKGCKYFSDGTFVCEGESDPRNPACTLNCDPRIAYGVCPRSEK 56

RESULT 9
US-08-454-295-6
; Sequence 6, Application US/08454295
; Patent No. 6031087
```

```
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,295
; FILING DATE: 01-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 58 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-454-295-6
Query Match      23.9%; Score 204; DB 3; Length 58;
Best Local Similarity 63.0%; Pred. No. 4.8e-12;
Matches 34; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 53 ICTNCCSGYKGNYSAGDLICEGSDPRNPDKCTFECDTQIAYSKCPRSEK 106
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 3 ICTNCCAGTKGCKYFSDGTFVCEGESDPRNPACTLNCDPRIAYGVCPRSEK 56

RESULT 10
US-09-431-500A-6
; Sequence 6, Application US/09431500A
; Patent No. 6261821
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; FILE REFERENCE: 9748B
; CURRENT APPLICATION NUMBER: US/09/431,500A
; CURRENT FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 08/454,295
; PRIOR FILING DATE: 1995-09-01
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Nicotiana glauca
US-09-431-500A-6
```

```
Query Match      23.9%; Score 204; DB 3; Length 58;
Best Local Similarity 63.0%; Pred. No. 4.8e-12;
Matches 34; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 53 ICTNCCSGYKGNYSKAGDLICEGSDPRNPKDCTFECDTQIAYSKCPRSEK 106
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3 ICTNCCAGTKGCKYFSDGDTFVCEGSDPRNPKACPRNCDPRIAYGICPLAEK 56

RESULT 11
US-09-431-498-6
; Sequence 6, Application US/09431498
; Patent No. 6440727
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/431,498
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/454,295
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 58 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-431-498-6

Query Match      23.9%; Score 204; DB 4; Length 58;
Best Local Similarity 63.0%; Pred. No. 4.8e-12;
Matches 34; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 53 ICTNCCSGYKGNYSKAGDLICEGSDPRNPKDCTFECDTQIAYSKCPRSEK 106
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3 ICTNCCAGTKGCKYFSDGDTFVCEGSDPRNPKACPRNCDPRIAYGICPLAEK 56

RESULT 12
US-09-431-499-6
; Sequence 6, Application US/09431499
; Patent No. 6451573
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
```

```
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF
; TITLE OF INVENTION: AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/431,499
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,295
; FILING DATE: 01-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 58 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-431-499-6

Query Match      23.9%; Score 204; DB 4; Length 58;
Best Local Similarity 63.0%; Pred. No. 4.8e-12;
Matches 34; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 53 ICTNCCSGYKGNYSKAGDLICEGSDPRNPKDCTFECDTQIAYSKCPRSEK 106
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3 ICTNCCAGTKGCKYFSDGDTFVCEGSDPRNPKACPRNCDPRIAYGICPLAEK 56

RESULT 13
US-08-454-295-7
; Sequence 7, Application US/08454295
; Patent No. 6031087
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

```
/ APPLICATION NUMBER: US/08/454,295
/ FILING DATE: 01-SEP-1995
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Digiglio, Frank S.
/ REGISTRATION NUMBER: 31,346
/ REFERENCE/DOCKET NUMBER: 9748
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (516) 742-4343
/ TELEFAX: (516) 742-4366
/ TELEX: 230 901 SANS UR
/ INFORMATION FOR SEQ ID NO: 7:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 58 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-454-295-7

Query Match      23.7%; Score 203; DB 3; Length 58;
Best Local Similarity 63.0%; Pred. No. 6e-12;
Matches 34; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 53 ICTNCCSGYKGNYSAGKDLICEGSDPRNPKDCTFECDTQIAYSKCPRSEK 106
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 3 ICTNCCAGKKGCKYFSDGTFVCEGESDPKNPKACPRNCDGRIAYGICPLSEK 56

RESULT 14
US-08-454-295-8
/ Sequence 8, Application US/08454295
/ Patent No. 6031087
/ GENERAL INFORMATION:
/ APPLICANT: Anderson, Marilyn A.
/ APPLICANT: Atkinson, Angela H.
/ APPLICANT: Heath, Robyn L.
/ APPLICANT: Clarke, Adrienne E.
/ TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
/ NUMBER OF SEQUENCES: 14
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Scully, Scott, Murphy & Presser
/ STREET: 400 Garden City Plaza
/ CITY: Garden City
/ STATE: New York
/ COUNTRY: United States of America
/ ZIP: 11530
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/454,295
/ FILING DATE: 01-SEP-1995
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Digiglio, Frank S.
/ REGISTRATION NUMBER: 31,346
/ REFERENCE/DOCKET NUMBER: 9748
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (516) 742-4343
/ TELEFAX: (516) 742-4366
/ TELEX: 230 901 SANS UR
/ INFORMATION FOR SEQ ID NO: 8:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 58 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-454-295-8
```

```
Query Match      23.7%; Score 203; DB 3; Length 58;
Best Local Similarity 63.0%; Pred. No. 6e-12;
Matches 34; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 53 ICTNCCSGYKGNYSAGKDLICEGSDPRNPKDCTFECDTQIAYSKCPRSEK 106
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 3 ICTNCCAGKKGCKYFSDGTFVCEGESDPKNPKACPRNCDGRIAYGICPLSEK 56

RESULT 15
US-09-431-500A-7
/ Sequence 7, Application US/09431500A
/ Patent No. 6261821
/ GENERAL INFORMATION:
/ APPLICANT: Anderson, Marilyn A.
/ APPLICANT: Atkinson, Angela H.
/ APPLICANT: Heath, Robyn L.
/ APPLICANT: Clarke, Adrienne E.
/ TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
/ FILE REFERENCE: 9748B
/ CURRENT APPLICATION NUMBER: US/09/431,500A
/ PRIOR FILING DATE: 1999-11-01
/ PRIOR FILING DATE: 1995-09-01
/ NUMBER OF SEQ ID NOS: 16
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 7
/ LENGTH: 58
/ TYPE: PRT
/ ORGANISM: Nicotiana glauca
/ US-09-431-500A-7

Query Match      23.7%; Score 203; DB 3; Length 58;
Best Local Similarity 63.0%; Pred. No. 6e-12;
Matches 34; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 53 ICTNCCSGYKGNYSAGKDLICEGSDPRNPKDCTFECDTQIAYSKCPRSEK 106
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 3 ICTNCCAGKKGCKYFSDGTFVCEGESDPKNPKACPRNCDGRIAYGICPLSEK 56

Search completed: January 24, 2005, 16:53:33
Job time : 41 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2005, 16:47:41 ; Search time 55 Seconds  
(without alignments)  
972.197 Million cell updates/sec

Title: US-10-725-829-2  
Perfect score: 855  
Sequence: 1 MAVHKVSFLACLLVLGMFL.....QDGFVCEGSEPEPKTTAYF 148

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCT05\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	855	100.0	148	17	US-10-725-829-2
2	615	71.9	152	17	US-10-725-829-4
3	379	44.3	368	10	US-09-812-502-3
4	379	44.3	368	14	US-10-164-961-3
5	379	44.3	368	14	US-10-157-622-3
6	379	44.3	368	14	US-10-072-809A-57
7	224	26.2	58	10	US-09-812-502-5
8	224	26.2	58	14	US-10-164-961-5
9	224	26.2	58	14	US-10-157-622-5
10	204	23.9	58	10	US-09-812-502-6
11	204	23.9	58	14	US-10-164-961-6
12	204	23.9	58	14	US-10-157-622-6
13	203	23.7	58	10	US-09-812-502-7

14	203	23.7	58	10	US-09-812-502-8	Sequence 8, Appl1
15	203	23.7	58	14	US-10-164-961-7	Sequence 7, Appl1
16	203	23.7	58	14	US-10-164-961-8	Sequence 8, Appl1
17	203	23.7	58	14	US-10-157-622-7	Sequence 7, Appl1
18	203	23.7	58	14	US-10-157-622-8	Sequence 8, Appl1
19	199	23.3	58	10	US-09-812-502-9	Sequence 9, Appl1
20	199	23.3	58	14	US-10-164-961-9	Sequence 9, Appl1
21	199	23.3	58	14	US-10-157-622-9	Sequence 9, Appl1
22	127	14.9	50	10	US-09-991-209-13	Sequence 13, Appl1
23	120	14.0	54	10	US-09-812-502-10	Sequence 10, Appl1
24	120	14.0	54	14	US-10-164-961-10	Sequence 10, Appl1
25	120	14.0	54	14	US-10-157-622-10	Sequence 10, Appl1
26	101.5	11.9	78	17	US-10-425-115-349792	Sequence 349792
27	100.5	11.8	1057	14	US-10-189-971-6	Sequence 6, Appl1
28	100.5	11.8	1192	14	US-10-189-971-18	Sequence 18, Appl1
29	100.5	11.8	1207	14	US-10-189-971-20	Sequence 20, Appl1
30	100.5	11.8	1251	14	US-10-189-971-16	Sequence 16, Appl1
31	100.5	11.8	1342	14	US-10-189-971-24	Sequence 24, Appl1
32	100.5	11.8	1477	14	US-10-189-971-8	Sequence 8, Appl1
33	100.5	11.8	1512	14	US-10-189-971-10	Sequence 10, Appl1
34	100.5	11.8	1535	14	US-10-189-971-14	Sequence 14, Appl1
35	100.5	11.8	1570	14	US-10-189-971-12	Sequence 12, Appl1
36	100.5	11.8	1593	14	US-10-189-971-4	Sequence 4, Appl1
37	100.5	11.8	1617	14	US-10-184-644-235	Sequence 235, App
38	100.5	11.8	1617	14	US-10-184-634-235	Sequence 235, App
39	100.5	11.8	1617	14	US-10-063-685-61	Sequence 61, Appl1
40	100.5	11.8	1628	14	US-10-189-971-2	Sequence 2, Appl1
41	99	11.6	214	16	US-10-767-701-36663	Sequence 36663, A
42	98	11.5	79	17	US-10-425-115-349791	Sequence 349791,
43	96.5	11.3	1574	10	US-09-825-751A-77	Sequence 77, Appl1
44	95.5	11.2	578	9	US-09-908-322-13	Sequence 13, Appl1
45	95.5	11.2	578	10	US-09-783-931-13	Sequence 13, Appl1

ALIGNMENTS

RESULT 1  
US-10-725-829-2  
; Sequence 2, Application US/10725829  
; Publication No. US20040205846A1  
; GENERAL INFORMATION:  
; APPLICANT: Chye, Nee Len  
; APPLICANT: Xu, Zeng-Fu  
; APPLICANT: Sin, Suk Fong  
; TITLE OF INVENTION: Genetically Modified Plants with Heterogenous Proteinase Inhibitor  
; TITLE OF INVENTION: SAPI2A or SAPI2B and Methods of Use Thereof for the Inhibition  
; TITLE OF INVENTION: of Trypsin- and Chymotrypsin-Like Activities  
; FILE REFERENCE: 9661-043-999  
; CURRENT APPLICATION NUMBER: US/10/725,829  
; CURRENT FILING DATE: 2003-12-01  
; PRIOR APPLICATION NUMBER: 60/429,992  
; PRIOR FILING DATE: 2002-11-29  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2  
; LENGTH: 148  
; TYPE: PRT  
; ORGANISM: Solanum americanum  
US-10-725-829-2

Query Match	100.0%;	Score 855;	DB 17;	Length 148;	
Best Local Similarity	100.0%;	Pred. No. 1.8e-71;			
Matches 148;	Conservative	0;	Mismatches	0;	
		Indels	0;	Gaps	0;
Qy	1	MAVHKVSFLACLLVLGMFLAKHVDKACTREC	GHFSYGCIPRSGSPQKPICTNCCSG	60	
Db	1	MAVHKVSFLACLLVLGMFLAKHVDKACTREC	GHFSYGCIPRSGSPQKPICTNCCSG	60	
Qy	61	YKGCNYSAGDGLICGESDPRNPDKCTPEC	DTQIAYSKCPRSEGKMIKPTGCTTCCTG	120	
Db	61	YKGCNYSAGDGLICGESDPRNPDKCTPEC	DTQIAYSKCPRSEGKMIKPTGCTTCCTG	120	

QY 121 YQCYFYDQDGFVCEGESPEPKTYAF 148  
Db 121 YQCYFYDQDGFVCEGESPEPKTYAF 148

RESULT 2  
US-10-725-829-4  
; Sequence 4, Application US/10725829  
; Publication No. US20040205846A1  
; GENERAL INFORMATION:  
; APPLICANT: Chye, Mee Len  
; APPLICANT: Xu, Zeng-Fu  
; APPLICANT: Sin, Suk Fong  
; TITLE OF INVENTION: Genetically Modified Plants with Heterogenous Proteinase Inhibitor  
; TITLE OF INVENTION: SAPIN2A or SAPIN2B and Methods of Use Thereof for the Inhibition  
; TITLE OF INVENTION: of Trypsin- and Chymotrypsin-like Activities  
; FILE REFERENCE: 9661-043-999  
; CURRENT APPLICATION NUMBER: US/10/725,829  
; CURRENT FILING DATE: 2003-12-01  
; PRIOR FILING DATE: 2002-11-29  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 4  
; LENGTH: 152  
; TYPE: PRT  
; ORGANISM: Solanum americanum  
US-10-725-829-4

Query Match 71.9%; Score 615; DB 17; Length 152;  
Best Local Similarity 75.3%; Pred. No. 3.3e-49;  
Matches 110; Conservative 8; Mismatches 24; Indels 4; Gaps 2;

QY 1 MAVHK-VSFLACILVLGMFL---LAKGVDAKACTRECGHFSYGI CPRSGSPQKPICTN 56  
Db 1 MAVHKEVSSLAYLLVLGLMFLHVSXVHKVDAKPTRECGNLGYGICPRSGSPENPICTN 60

QY 57 CCSGYGKCNYYSAKGDLICGESDPPRNPXDCCTPECDTQIAYSKCPRSEGKMIKPTGCTT 116  
Db 61 CCSGYGKCNYYSAANGFICGSSDPKPNPTCLFCDGDIAYSKCPRSEGTIIYPTGCTT 120

QY 117 CTTGYGQCYFYDQDGFVCEGESPEP 142  
Db 121 CTTGYGQCYFYFSKEGEFVCEGESDEP 146

RESULT 3  
US-09-812-502-3  
; Sequence 3, Application US/09812502  
; Publication No. US20030027303A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, Marilyn A.  
; APPLICANT: Atkinson, Angela H.  
; APPLICANT: Heath, Robyn L.  
; APPLICANT: Clarke, Adrienne E.  
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC  
; TITLE OF INVENTION: SEQUENCES ENCODING SAME  
; FILE REFERENCE: 9748B  
; CURRENT APPLICATION NUMBER: US/09/812,502  
; CURRENT FILING DATE: 2001-03-20  
; PRIOR APPLICATION NUMBER: US/09/431,500  
; PRIOR FILING DATE: 1999-11-01  
; PRIOR APPLICATION NUMBER: 08/454,295  
; PRIOR FILING DATE: 1995-09-01  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 368  
; TYPE: PRT  
; ORGANISM: Nicotiana glauca  
US-09-812-502-3

Query Match 44.3%; Score 379; DB 10; Length 368;

Best Local Similarity 58.1%; Pred. No. 5.8e-27;  
Matches 68; Conservative 9; Mismatches 36; Indels 4; Gaps 3;

QY 28 KACTRECG-HFSYGI CPRSGSPQKPICTNCCSGYKCNYYSAKGLICGESDPPRNPX 86  
Db 1 KACTLNCDPRIAYGVCPRSEKKNDRICNCCAGTGGCKYFSDGTFVCEGESDPPRNPKA 60

QY 87 CTFECDTQIAYSKCPRSEGKMIKPTGCTTCTTGQCYFYDQDGFVCEGESPEPK 143  
Db 61 CTILNCDPRIAYGVCPRSEKK--NDRICNCCAGTGGCKYFSDGTFVCEGES--DPR 114

RESULT 4  
US-10-164-961-3  
; Sequence 3, Application US/10164961  
; Publication No. US200310096388A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, Marilyn A.  
; APPLICANT: Atkinson, Angela H.  
; APPLICANT: Heath, Robyn L.  
; APPLICANT: Clarke, Adrienne E.  
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scully, Scott, Murphy & Presser  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/164,961  
; FILING DATE: 07-Jun-2002  
; CLASSIFICATION: <Unknown>  
; APPLICATION DATA:  
; APPLICATION NUMBER: US/08/454,295  
; FILING DATE: 01-SEP-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DiGiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 9748  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 368 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-10-164-961-3

Query Match 44.3%; Score 379; DB 14; Length 368;  
Best Local Similarity 58.1%; Pred. No. 5.8e-27;  
Matches 68; Conservative 9; Mismatches 36; Indels 4; Gaps 3;

QY 28 KACTRECG-HFSYGI CPRSGSPQKPICTNCCSGYKCNYYSAKGLICGESDPPRNPX 86  
Db 1 KACTLNCDPRIAYGVCPRSEKKNDRICNCCAGTGGCKYFSDGTFVCEGESDPPRNPKA 60

QY 87 CTFECDTQIAYSKCPRSEGKMIKPTGCTTCTTGQCYFYDQDGFVCEGESPEPK 143  
Db 61 CTILNCDPRIAYGVCPRSEKK--NDRICNCCAGTGGCKYFSDGTFVCEGES--DPR 114

RESULT 5

US-10-157-622-3  
; Sequence 3, Application US/10157622  
; Publication No. US20030129720A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, Marilyn A.  
; Atkinson, Angela H.  
; Heath, Robyn L.  
; Clarke, Adrienne E.  
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scully, Scott, Murphy & Presser  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/157,622  
; FILING DATE: 29-May-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/454,295  
; FILING DATE: 01-SEP-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DiGiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 9748  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 368 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-10-157-622-3  
Query Match 44.3%; Score 379; DB 14; Length 368;  
Best Local Similarity 58.1%; Pred. No. 5.8e-27;  
Matches 68; Conservative 9; Mismatches 36; Indels 4; Gaps 3;  
QY 28 KACTREGC-HFSGICPRSEGSQPKICTNCCSGYKGNYSAGDGLICGSDPRNPKD 86  
DB 1 KACTLNCDPRIAYGVCPRSEKKNDRICNCCAGTKGCKYFSDGTFCVCEGSDPRNPKA 60  
QY 87 CTPECDTQIAYSKPRSEGMIIKPTGCTTCCTGYGCGYFFDQDGFVCEGSESPK 143  
DB 61 CTLNCDPRIAYGVCPRSEKK--NDRICNCCAGTKGCKYFSDGTFCVCEGSDPR 114  
RESULT 6  
US-10-072-809A-57  
; Sequence 57, Application US/10072809A  
; Publication No. US20030217382A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, Marilyn A., Lay, Fung T., Heath, Robyn, L.  
; TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and u  
; FILE OF INVENTION: therefor  
; FILE REFERENCE: 18-01  
; CURRENT APPLICATION NUMBER: US/10/072,809A  
; PRIOR FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: USSN 60/267,271  
; PRIOR FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 61  
US-10-157-622-3  
Query Match 44.3%; Score 379; DB 14; Length 368;  
Best Local Similarity 58.1%; Pred. No. 5.8e-27;  
Matches 68; Conservative 9; Mismatches 36; Indels 4; Gaps 3;  
QY 28 KACTREGC-HFSGICPRSEGSQPKICTNCCSGYKGNYSAGDGLICGSDPRNPKD 86  
DB 1 KACTLNCDPRIAYGVCPRSEKKNDRICNCCAGTKGCKYFSDGTFCVCEGSDPRNPKA 60  
QY 87 CTPECDTQIAYSKPRSEGMIIKPTGCTTCCTGYGCGYFFDQDGFVCEGSESPK 143  
DB 61 CTLNCDPRIAYGVCPRSEKK--NDRICNCCAGTKGCKYFSDGTFCVCEGSDPR 114  
RESULT 6  
US-10-072-809A-57  
; Sequence 57, Application US/10072809A  
; Publication No. US20030217382A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, Marilyn A., Lay, Fung T., Heath, Robyn, L.  
; TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and u  
; FILE OF INVENTION: therefor  
; FILE REFERENCE: 18-01  
; CURRENT APPLICATION NUMBER: US/10/072,809A  
; PRIOR FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: USSN 60/267,271  
; PRIOR FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 61

; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 57  
; LENGTH: 368  
; TYPE: PRT  
; ORGANISM: Nicotiana glauca  
US-10-072-809A-57  
Query Match 44.3%; Score 379; DB 14; Length 368;  
Best Local Similarity 58.1%; Pred. No. 5.8e-27;  
Matches 68; Conservative 9; Mismatches 36; Indels 4; Gaps 3;  
QY 28 KACTREGC-HFSGICPRSEGSQPKICTNCCSGYKGNYSAGDGLICGSDPRNPKD 86  
DB 1 KACTLNCDPRIAYGVCPRSEKKNDRICNCCAGTKGCKYFSDGTFCVCEGSDPRNPKA 60  
QY 87 CTPECDTQIAYSKPRSEGMIIKPTGCTTCCTGYGCGYFFDQDGFVCEGSESPK 143  
DB 61 CTLNCDPRIAYGVCPRSEKK--NDRICNCCAGTKGCKYFSDGTFCVCEGSDPR 114  
RESULT 7  
US-09-812-502-5  
; Sequence 5, Application US/09812502  
; Publication No. US20030027303A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, Marilyn A.  
; APPLICANT: Atkinson, Angela H.  
; APPLICANT: Heath, Robyn L.  
; APPLICANT: Clarke, Adrienne E.  
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC  
; TITLE OF INVENTION: SEQUENCES ENCODING SAME  
; FILE REFERENCE: 9748B  
; CURRENT APPLICATION NUMBER: US/09/812,502  
; CURRENT FILING DATE: 2001-03-20  
; PRIOR APPLICATION NUMBER: US/09/431,500  
; PRIOR FILING DATE: 1999-11-01  
; PRIOR APPLICATION NUMBER: 08/454,295  
; PRIOR FILING DATE: 1995-09-01  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 58  
; TYPE: PRT  
; ORGANISM: Nicotiana glauca  
US-09-812-502-5  
Query Match 26.2%; Score 224; DB 10; Length 58;  
Best Local Similarity 68.5%; Pred. No. 2.1e-13;  
Matches 37; Conservative 4; Mismatches 13; Indels 0; Gaps 0;  
QY 53 ICTNCCSGYKGNYSAGDGLICGSDPRNPKDCTPECDTQIAYSKCPRSEK 106  
DB 3 ICTNCCAGTKGCKYFSDGTFCVCEGSDPRNPKACTLNCDPRIAYGVCPRSEK 56  
RESULT 8  
US-10-164-961-5  
; Sequence 5, Application US/10164961  
; Publication No. US20030096388A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, Marilyn A.  
; Atkinson, Angela H.  
; Heath, Robyn L.  
; Clarke, Adrienne E.  
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scully, Scott, Murphy & Presser  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11530

```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/164,961
; FILING DATE: 07-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,295
; FILING DATE: 01-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 58 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; MOLECULE TYPE: protein
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-164-961-5
;
; Query Match 26.2%; Score 224; DB 14; Length 58;
; Best Local Similarity 68.5%; Pred. No. 2.1e-13;
; Matches 37; Conservative 4; Mismatches 13; Indels 0; Gaps 0;
;
; QY 53 ICTNCCSGYKGNYSAGKGLICEGSDPRNPKDCTFECDTQIAYSKCPRSEK 106
; DB 3 ICTNCCAGTKGCKYFSDGTFVCEGESDPRNPKACTLNCDPRIAYGVCPRSEK 56
;
; RESULT 9
; US-10-157-622-5
; Sequence 5, Application US/10157622
; Publication No. US20030129720A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; Heath, Robyn L.
; Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/157,622
; FILING DATE: 29-May-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,295
; FILING DATE: 01-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
;
; US-10-164-961-5
;
; Query Match 26.2%; Score 224; DB 14; Length 58;
; Best Local Similarity 68.5%; Pred. No. 2.1e-13;
; Matches 37; Conservative 4; Mismatches 13; Indels 0; Gaps 0;
;
; QY 53 ICTNCCSGYKGNYSAGKGLICEGSDPRNPKDCTFECDTQIAYSKCPRSEK 106
; DB 3 ICTNCCAGTKGCKYFSDGTFVCEGESDPRNPKACTLNCDPRIAYGVCPRSEK 56
;
; RESULT 10
; US-09-812-502-6
; Sequence 6, Application US/09812502
; Publication No. US20030027303A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; FILE REFERENCE: 9748B
; CURRENT APPLICATION NUMBER: US/09/812,502
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US/09/431,500
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 08/454,295
; PRIOR FILING DATE: 1995-09-01
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Nicotiana glauca
; US-09-812-502-6
;
; Query Match 23.9%; Score 204; DB 10; Length 58;
; Best Local Similarity 63.0%; Pred. No. 1.5e-11;
; Matches 34; Conservative 5; Mismatches 15; Indels 0; Gaps 0;
;
; QY 53 ICTNCCSGYKGNYSAGKGLICEGSDPRNPKDCTFECDTQIAYSKCPRSEK 106
; DB 3 ICTNCCAGTKGCKYFSDGTFVCEGESDPRNPKACTLNCDPRIAYGVCPRSEK 56
;
; RESULT 11
; US-10-164-961-6
; Sequence 6, Application US/10164961
; Publication No. US20030096388A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; Atkinson, Angela H.
; Heath, Robyn L.
; Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
```

```
STATE: New York
COUNTRY: United States of America
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/164,961
FILING DATE: 07-Jun-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/454,295
FILING DATE: 01-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: DiGiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9748
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 58 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-164-961-6

Query Match 23.9%; Score 204; DB 14; Length 58;
Best Local Similarity 63.0%; Pred. No. 1.5e-11;
Matches 34; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 53 ICTNCCSGYKGCNYYSAGKDLICEGSDPRNPKDCTFECDTQIAYSKCPRSEK 106
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3 ICTNCCAGTKGCKYFSDGTFVCEGSDPRNPKACPRNCDPRIAYGICPLAEK 56

RESULT 13
US-09-812-502-7
Sequence 7, Application US/09812502
Publication No. US20030027303A1
GENERAL INFORMATION:
APPLICANT: Anderson, Marilyn A.
APPLICANT: Atkinson, Angela H.
APPLICANT: Heath, Robyn L.
APPLICANT: Clarke, Adrienne E.
TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
TITLE OF INVENTION: SEQUENCES ENCODING SAME
FILE REFERENCE: 9748B
CURRENT APPLICATION NUMBER: US/09/812,502
PRIOR FILING DATE: 2001-03-20
PRIOR APPLICATION NUMBER: US/09/431,500
PRIOR FILING DATE: 1999-11-01
PRIOR APPLICATION NUMBER: 08/454,295
PRIOR FILING DATE: 1995-09-01
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 58
TYPE: PRT
ORGANISM: Nicotiana glauca
US-09-812-502-7

Query Match 23.7%; Score 203; DB 10; Length 58;
Best Local Similarity 63.0%; Pred. No. 1.9e-11;
Matches 34; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 53 ICTNCCSGYKGCNYYSAGKDLICEGSDPRNPKDCTFECDTQIAYSKCPRSEK 106
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3 ICTNCCAGTKGCKYFSDGTFVCEGSDPRNPKACPRNCDPRIAYGICPLAEK 56

RESULT 14
US-09-812-502-8
Sequence 8, Application US/09812502
Publication No. US20030027303A1
GENERAL INFORMATION:
APPLICANT: Anderson, Marilyn A.
APPLICANT: Atkinson, Angela H.
APPLICANT: Heath, Robyn L.
APPLICANT: Clarke, Adrienne E.
TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
TITLE OF INVENTION: SEQUENCES ENCODING SAME
FILE REFERENCE: 9748B
```

```

Matches 34; Conservative 5; Mismatches 15; Indels 0; Gaps 0;
Qy 53 ICTCCSGYKGNYSYSAKGLTCEGSDPRNPKDCTFECDTQIAYSKCPRSEK 106
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3 ICTCCAGKKGCKYFDDGTFVCEGSDPRNPKACPRNCDGR IAYIGCLPSEK 56

Search completed: January 24, 2005, 16:54:34
Job time : 56 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2005, 16:41:10 ; Search time 67 Seconds  
(without alignments)  
792.417 Million cell updates/sec

Title: US-10-725-829-2  
Perfect score: 855  
Sequence: 1 MAVHKVSFLACLLVLGWF.....QDGFVCEGSEBPCKTAYF 148

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001s:\*
- 5: Geneseq2002s:\*
- 6: Geneseq2003as:\*
- 7: Geneseq2003bs:\*
- 8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	855	100.0	148	ADP02855	Adp02855 Solanum a
2	615	71.9	152	ADP02857	Adp02857 Solanum a
3	585.5	68.5	153	AAR20099	Aar20099 Wound-ind
4	379	44.3	368	AAR54135	Aar54135 Deduced s
5	379	44.3	368	ADI56864	Adi56864 Plant def
6	224	26.2	58	AAR54125	Aar54125 Sequence
7	206	24.1	58	AAR54129	Aar54129 Sequence
8	204	23.9	58	AAR54126	Aar54126 Sequence
9	203	23.7	58	AAR54127	Aar54127 Sequence
10	203	23.7	58	AAR54128	Aar54128 Sequence
11	127	14.9	50	ADG69883	Adg69883 Potato pr
12	118	13.6	54	AAR54130	Aar54130 Sequence
13	116.5	13.6	49	ADN88896	Adn88896 Potato pr
14	106	12.4	444	ADP31478	Adp31478 Human sec
15	104.5	12.2	186	ADP30830	Adp30830 Human sec
16	104	12.2	137	ABM74314	Abm74314 DNA clone
17	100.5	11.8	1057	6 ABP97370	Abp97370 Human kie
18	100.5	11.8	1192	6 ABP97376	Abp97376 Human kie
19	100.5	11.8	1207	6 ABP97377	Abp97377 Human kie
20	100.5	11.8	1251	6 ABP97375	Abp97375 Human kie
21	100.5	11.8	1342	6 ABP97379	Abp97379 Human kie
22	100.5	11.8	1477	6 ABP97371	Abp97371 Human kie
23	100.5	11.8	1512	6 ABP97372	Abp97372 Human kie
24	100.5	11.8	1535	6 ABP97374	Abp97374 Human kie
25	100.5	11.8	1570	6 ABP97373	Abp97373 Human kie

26	100.5	11.8	1593	6 ABP97369	Abp97369 Human kie
27	100.5	11.8	1628	6 ABP97368	Abp97368 Human kie
28	100	11.7	3046	8 ADP31684	Adp31684 Human sec
29	97	11.3	434	4 ABB66756	Abb66756 Drosophil
30	96	11.2	623	5 ABB77796	Abb77796 Amino aci
31	96	11.2	1050	5 ABB77795	Abb77795 Antigen e
32	96	11.2	2108	8 ADN96130	Adn96130 Human NOV
33	95.5	11.2	193	2 AAY08491	Aay08491 Rat serin
34	95.5	11.2	216	2 AAY08492	Aay08492 Rat serin
35	95.5	11.2	405	5 ABB77797	Abb77797 Amino aci
36	95.5	11.2	1359	8 ADP31596	Adp31596 Human sec
37	95	11.1	585	8 ADP31445	Adp31445 Human sec
38	95	11.1	917	8 ADP30890	Adp30890 Human sec
39	94	11.0	148	2 AAY09280	Aay09280 YL-4 ther
40	93.5	10.9	1812	8 ADP31556	Adp31556 Human sec
41	93.5	10.9	1911	8 ADP31555	Adp31555 Human sec
42	92.5	10.8	2321	2 AAW49698	Aaw49698 Human Not
43	92.5	10.8	2321	7 ADJ69828	Adj69828 Human hea
44	92.5	10.8	2321	8 ADH34621	Adh34621 Notch hom
45	92.5	10.8	2321	8 ADJ75570	Adj75570 Marker ge

## ALIGNMENTS

## RESULT 1

## ADP02855

ID ADP02855 standard; protein; 148 AA.

XX AC ADP02855;

XX DT 09-SEP-2004 (first entry)

XX DE Solanum americanum proteinase inhibitor II protein SAPIN2a.

XX KW proteinase inhibitor II; transformed plant; resistance; insect; pest;

XX KW pathogen; programmed cell death; senescence.

XX OS Solanum americanum.

XX PN WO2004050873-A1.

XX PD 17-JUN-2004.

XX PF 01-DEC-2003; 2003WO-CN001020.

XX PR 29-NOV-2002; 2002US-0429992P.

XX PA (UYHK-) UNIV HONG KONG.

XX PI Chye M, Xu Z, Sin S;

XX DR WPI; 2004-450731/42.

XX DR N-PSDB; ADP02854.

PT New proteinase inhibitor, SAPIN2a or SAPIN2b, useful in producing transformed plants having enhanced resistance to insects, pest or pathogens and in which programmed cell death or senescence is inhibited.

XX PS Claim 5; SEQ ID NO 2; 90pp; English.

XX CC The invention relates to an isolated proteinase inhibitor II nucleic acid molecule (I). The proteinase inhibitor II nucleic acid molecule and the encoded polypeptide, methods are useful in producing transgenic plants having enhanced resistance to insects, pest or pathogens and in which programmed cell death or senescence is inhibited. This sequence corresponds to the American black nightshade SAPIN2a protein.

XX SQ Sequence 148 AA;

Query Match 100.0%; Score 855; DB 8; Length 148;

Best Local Similarity 100.0%; Pred. No. 36-64;

Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;





AD AARS4135 standard; protein; 368 AA.  
 AC AARS4135;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 02-MAR-1995 (first entry)  
 XX  
 DE Deduced sequence of mature N-alata PI precursor.  
 XX  
 KW Type II serine proteinase inhibitor precursor; PI; tobacco;  
 KW transgenic plant; anti-pathogen; anti-predator.  
 KW Nicotiana alata.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 1..58  
 FT /label= 1  
 FT Peptide 1..24  
 FT /label= 1  
 FT Active-site 5..6  
 FT /label= reactive site  
 FT Peptide 25..82  
 FT /label= 2  
 FT Domain 59..116  
 FT /label= 2  
 FT Active-site 63..64  
 FT /label= reactive site  
 FT Peptide 83..140  
 FT /label= 3  
 FT Domain 117..174  
 FT /label= 3  
 FT Active-site 121..122  
 FT /label= reactive site  
 FT Peptide 141..198  
 FT /label= 4  
 FT Domain 175..232  
 FT /label= 4  
 FT Active-site 179..180  
 FT /label= reactive site  
 FT Peptide 199..256  
 FT /label= 5  
 FT Domain 233..290  
 FT /label= 5  
 FT Active-site 237..238  
 FT /label= reactive site  
 FT Peptide 257..314  
 FT /label= 6  
 FT Domain 291..343  
 FT /label= 6  
 FT Active-site 295..296  
 FT /label= reactive site  
 FT Peptide 315..368  
 FT /label= 7  
 XX  
 PN W09413810-A1.  
 XX  
 PD 23-JUN-1994.  
 XX  
 PF 16-DEC-1993; 93WO-AU000659.  
 XX  
 PF 16-DEC-1992; 92AU-00006399.  
 XX  
 PF (UYME ) UNIV MELBOURNE.  
 XX  
 PI Anderson MA, Atkinson AH, Heath RL, Clarke AE;  
 XX  
 DR N-PSDB; AARQ68728, AARQ68729.  
 DR WPI; 1994-217886/26.  
 XX  
 PT Nicotiana alata type II serine protease inhibitor precursor and DNA -  
 PT useful in prodn of anti-pathogen or anti-predator constructs for plants.  
 XX  
 PS Claim 16; Page 45-47; 83pp; English.

XX A cDNA library, prepd. from mRNA from the stigmas and styles of mature  
 CC flowers of N. alata was screened for clones of highly expressed genes  
 CC which were not associated with self-incompatibility genotype. Clones  
 CC encoding a protein with some identity to the type II proteinase  
 CC inhibitors from potato and tomato were selected. The largest clone, NA-PI  
 CC -2, is given in AARQ68729. The predicted AA sequence in AARS4135, AARQ68728  
 CC is the coding region of AARQ68729. The type II serine PI has six conserved  
 CC domains wherein the first two are 100% identical and contain chymotrypsin  
 CC -specific sites. The 3rd, 4th and 5th domains share 95-98% identity and  
 CC have sites specific for trypsin. The 6th domain also has a trypsin  
 CC specific site but less identity to the 3rd, 4th and 5th domains (79-90%)  
 CC due mainly to a divergent 3' sequence. It has a mol. wt. of approx. 42-  
 CC 45kDa with an approx. 29 AA signal sequence. The N-terminal sequence of  
 CC the monomeric PI is represented in each of the six repeated domains in  
 CC the predicted sequence of the PI precursor protein. Thus, it is likely  
 CC that the PI precursor protein is cleaved at six sites to produce seven  
 CC peptides. Six of the seven peptides, peptides 2-7 (AARS4125-R54130) would  
 CC be in the same mol. wt. range as the monomeric PI (about 6kDa) and would  
 CC have the same N-terminal sequence. Peptide 7 does not contain a consensus  
 CC site for trypsin or chymotrypsin. Peptide 1 (AARS4124) is small than  
 CC 6kDa, has a different N-terminus and was not detected in a purified  
 CC monomeric PI prepn. It could be envisaged that peptide 1 and peptide 7  
 CC would form a functional proenzyme inhibitor with the inhibitory site on  
 CC peptide 1 held in the correct conformation by disulphide bonds between  
 CC the two peptides. (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 368 AA;  
 Query Match 44.3%; Score 379; DB 2; Length 368;  
 Best Local Similarity 58.1%; Pred. No. 7.7e-24;  
 Matches 68; Conservative 9; Mismatches 36; Indels 4; Gaps 3;  
 QY 28 KACTREG-HFSYICPRSEGSQKPICTNCCSGYKGNYYSAKDLICGESDPRNPKD 86  
 DB 1 KACTLNCDPRIAYGVCPSESEKKNDRICTNCCAGTKGCKYFSDGTFCVCGESDPRNPKA 60  
 QY 87 CTPECDTQIAYSKCPRESEKMIKPTGCTTCCTGYOGCYFDDGDFVCGESPEPK 143  
 DB 61 CTLNCDPRIAYGVCPRESEKK--NDRICTNCCAGTKGCKYFSDGTFCVCGES--DPR 114  
 RESULT 5  
 ADI56864  
 ID ADI56864 standard; protein; 368 AA.  
 XX  
 AC ADI56864;  
 XX  
 DT 22-APR-2004 (first entry)  
 XX  
 DE Plant defensin NAPI mature domain.  
 DE  
 KW insecticide; antifungal; virucide; antibacterial; flower development;  
 KW plant pest resistance; floral defensin precursor; transgenic plant;  
 KW insect resistance.  
 XX  
 OS Unidentified.  
 XX  
 PN W0200263011-A1.  
 XX  
 PD 15-AUG-2002.  
 XX  
 PF 08-FEB-2002; 2002WO-AU000123.  
 XX  
 PF 08-FEB-2001; 2001US-0267271P.  
 XX  
 PA (HEXI-) HEXIMA LTD.  
 XX  
 PI Anderson MA, Lay FT, Heath RL;  
 XX  
 DR WPI; 2002-657538/70.  
 DR N-PSDB; ADI56863.  
 XX

PT New isolated plant floral defensin-like polypeptides and polynucleotides,  
 PT for generating transgenic plants having resistance or at least reduced  
 PT sensitivity to plant pests including insects, microorganisms, fungi  
 and/or viruses.

PS Disclosure; SEQ ID NO 57; 164pp; English.

XX The invention relates to an isolated polypeptide comprising, in its  
 CC precursor form, an N-terminal signal domain, a mature domain and an  
 CC acidic C-terminal domain, where the polypeptide is produced during flower  
 CC development and its mature domain has activity against one or more plant  
 CC pests. Plant floral defensin-like polypeptides and polynucleotides are  
 CC useful in generating transgenic plants having resistance or at least  
 CC reduced sensitivity to plant pests including insects, microorganisms,  
 CC fungi and/or viruses. They are also useful in generating recombinant  
 CC defensin-like molecules for use in the topical application of  
 CC compositions to prevent or retard pest-infestation of plants. The floral-  
 CC and seed-derived defensins are useful in the generation of insect  
 CC resistance in plants. This sequence represents the amino acid sequence of  
 CC the mature domain from a plant defensin protein.

XX Sequence 368 AA;

Query Match 44.3%; Score 379; DB 5; Length 368;  
 Best Local Similarity 58.1%; Pred. No. 7.7e-24;  
 Matches 68; Conservative 9; Mismatches 36; Indels 4; Gaps 3;

QY 28 KACTREG-HFSYGICPRSGSPQKPICTNCCSGYKGNYSAGDLCIGESDPRNPKD 86  
 Db 1 KACTLNCDPRIAYGVCPRSEBK--NDRICTNCCAGTCKYFSDGTFVCEGESDPRNPKA 60  
 QY 87 CTPECCTQIAYSKCPRSEGMIIKPTGCTTCCTGYGCGYVFDQDGFVCEGESPEPK 143  
 Db 61 CTLNCDPRIAYGVCPRSEBK--NDRICTNCCAGTCKYFSDGTFVCEGES-DPR 114

RESULT 6

AAAR54125  
 ID AAR54125 standard; peptide; 58 AA.

XX AAR54125;

XX 25-MAR-2003 (revised)

DT 02-MAR-1995 (first entry)

XX Sequence of peptide 2 of N-alata PI precursor.

XX Type II serine proteinase inhibitor precursor; PI; tobacco;  
 KW transgenic plant; anti-pathogen; anti-predator; peptide.

XX Nicotiana alata.

XX WO9413810-A1.

PN 23-JUN-1994.

XX 16-DEC-1993; 93WO-AU000659.

XX 16-DEC-1992; 92AU-00006399.

XX (UYME ) UNIV MELBOURNE.

XX Anderson MA, Atkinson AH, Heath RL, Clarke AE;

XX WPI; 1994-217886/26.

DR N-PSDB; AAQ68728, AAQ68729.

XX Nicotiana alata type II serine protease inhibitor precursor and DNA -  
 PT useful in prodn of anti-pathogen or anti-predator constructs for plants.

XX Claim 18; Page 50; 83pp; English.

XX A cDNA library, prepd. from mRNA from the stigmas and styles of mature

CC flowers of N. alata was screened for clones of highly expressed genes  
 CC which were not associated with self-incompatibility genotype. Clones  
 CC encoding a protein with some identity to the type II proteinase  
 CC inhibitors from potato and tomato were selected. The largest clone, NA-PI  
 CC -2, is given in AAQ68729. The predicted AA sequence in AAR54135, AAQ68728  
 CC domains wherein the first two are 100% identical and contain chymotrypsin  
 CC specific sites. The 3rd, 4th and 5th domains share 95-98% identity and  
 CC have sites specific for trypsin. The 6th domain also has a trypsin  
 CC specific site but less identity to the 3rd, 4th and 5th domains (79-90%)  
 CC due mainly to a divergent 3' sequence. It has a mol. st. of approx. 42-  
 CC 45kDa with an approx. 29 AA signal sequence. The N-terminal sequence of  
 CC the monomeric PI is represented in each of the six repeated domains in  
 CC the predicted sequence of the PI precursor protein. Thus, it is likely  
 CC that the PI precursor protein is cleaved at six sites to produce seven  
 CC peptides. Six of the seven peptides, peptides 2-7 (AAR54125-R54130) would  
 CC be in the same mol. wt. range as the monomeric PI (about 6kDa) and would  
 CC have the same N-terminal sequence. Peptide 7 does not contain a consensus  
 CC site for trypsin or chymotrypsin. Peptide 1 (AAR54124) is small than  
 CC 6kDa, has a different N-terminus and was not detected in a purified  
 CC monomeric PI prepn. It could be envisaged that peptide 1 and peptide 7  
 CC would form a functional proenzyme inhibitor with the inhibitory site on  
 CC peptide 1 held in the correct conformation by disulphide bonds between  
 CC the two peptides. A monomer which corresp. to any of the peptides 2-7 is  
 CC claimed. (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 58 AA;

Query Match 26.2%; Score 224; DB 2; Length 58;  
 Best Local Similarity 68.5%; Pred. No. 1.4e-11;

Matches 37; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 53 ICTNCCSGYKGNYSAGDLCIGESDPRNPKDCTFECDTQIAYSKCPRSECK 106

Db 3 ICTNCCAGTCKYFSDGTFVCEGESDPRNPKACTLNCDPRIAYGVCPRSEK 56

RESULT 7

AAAR54129

ID AAR54129 standard; peptide; 58 AA.

XX AAR54129;

XX 25-MAR-2003 (revised)

DT 02-MAR-1995 (first entry)

XX Sequence of peptide 6 of N-alata PI precursor.

XX Type II serine proteinase inhibitor precursor; PI; tobacco;  
 KW transgenic plant; anti-pathogen; anti-predator; peptide.

XX Nicotiana alata.

XX WO9413810-A1.

PN 23-JUN-1994.

XX 16-DEC-1993; 93WO-AU000659.

XX 16-DEC-1992; 92AU-00006399.

XX (UYME ) UNIV MELBOURNE.

XX Anderson MA, Atkinson AH, Heath RL, Clarke AE;

XX WPI; 1994-217886/26.

DR N-PSDB; AAQ68728, AAQ68729.

XX Nicotiana alata type II serine protease inhibitor precursor and DNA -  
 PT useful in prodn of anti-pathogen or anti-predator constructs for plants.

XX Claim 18; Page 52; 83pp; English.

XX



```

PS CC Claim 18; Page 51; 83pp; English.
XX CC A cDNA library, prepd. from mRNA from the stigmas and styles of mature
CC flowers of N. alata was screened for clones of highly expressed genes
CC which were not associated with self-incompatibility genotype. Clones
CC encoding a protein with some identity to the type II proteinase
CC inhibitors from potato and tomato were selected. The largest clone, NA-PI
CC -2, is given in AAQ68729. The predicted AA sequence in AAR54135. AAQ68728
CC is the coding region of AAQ68729. The type II serine PI has six conserved
CC domains wherein the first two are 100% identical and contain chymotrypsin
CC -specific sites. The 3rd, 4th and 5th domains share 95-98% identity and
CC have sites specific for trypsin. The 6th domain also has a trypsin
CC specific site but less identity to the 3rd, 4th and 5th domains (79-90%)
CC due mainly to a divergent 3' sequence. It has a mol. st. of approx. 42-
CC 45kDa with an approx. 29 AA signal sequence. The N-terminal sequence of
CC the monomeric PI is represented in each of the six repeated domains in
CC the predicted sequence of the PI precursor protein. Thus, it is likely
CC that the PI precursor protein is cleaved at six sites to produce seven
CC peptides. Six of the seven peptides, peptides 2-7 (AAR54125-R54130) would
CC be in the same mol. wt. range as the monomeric PI (about 6kDa) and would
CC have the same N-terminal sequence. Peptide 7 does not contain a consensus
CC site for trypsin or chymotrypsin. Peptide 1 (AAR54124) is small than
CC 6kDa, has a different N-terminus and was not detected in a purified
CC monomeric PI prep. It could be envisaged that peptide 1 and peptide 7
CC would form a functional proteinase inhibitor with the inhibitory site on
CC peptide 1 held in the correct conformation by disulphide bonds between
CC the two peptides. A monomer which corresp. to any of the peptides 2-7 is
CC claimed. (Updated on 25-MAR-2003 to correct PN field.)
XX CC
SQ Sequence 58 AA;

Query Match 23.7%; Score 203; DB 2; Length 58;
Best Local Similarity 63.0%; Pred. No. 8.2e-10;
Matches 34; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 53 ICTNCCSGYKGNYSAGKDLICEGESDPNPKDCTFECDTQIAYSKCPRSEK 106
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 3 ICTNCCAGKKGCKYFSDGTFVCEGESDPKNPKACPRNCDGRIAYGICPLSEK 56
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

RESULT 10
AAR54128
ID AAR54128 standard; peptide; 58 AA.
XX
AC AAR54128;
XX
DT 25-MAR-2003 (revised)
DT 02-MAR-1995 (first entry)
XX
DE Sequence of peptide 5 of N-alata PI precursor.
XX
KW Type II serine proteinase inhibitor precursor; PI; tobacco;
KW transgenic plant; anti-pathogen; anti-predator; peptide.
XX
OS Nicotiana alata.
XX
FN WO9413810-A1.
XX
PD 23-JUN-1994.
XX
PF 16-DEC-1993; 93WO-AU000659.
XX
PR 16-DEC-1992; 92AU-00006399.
XX
PA (UYME ) UNIV MELBOURNE.
XX
PI Anderson MA, Atkinson AH, Heath RL, Clarke AE;
XX
DR WPI; 1994-217886/26.
XX
DR N-PSDB; AAQ68728, AAQ68729.
XX
PT Nicotiana alata type II serine protease inhibitor precursor and DNA -
PT useful in prodn of anti-pathogen or anti-predator constructs for plants.

Claim 18; Page 52; 83pp; English.
XX CC A cDNA library, prepd. from mRNA from the stigmas and styles of mature
XX flowers of N. alata was screened for clones of highly expressed genes
XX which were not associated with self-incompatibility genotype. Clones
XX encoding a protein with some identity to the type II proteinase
XX inhibitors from potato and tomato were selected. The largest clone, NA-PI
XX -2, is given in AAQ68729. The predicted AA sequence in AAR54135. AAQ68728
XX is the coding region of AAQ68729. The type II serine PI has six conserved
XX domains wherein the first two are 100% identical and contain chymotrypsin
XX -specific sites. The 3rd, 4th and 5th domains share 95-98% identity and
XX have sites specific for trypsin. The 6th domain also has a trypsin
XX specific site but less identity to the 3rd, 4th and 5th domains (79-90%)
XX due mainly to a divergent 3' sequence. It has a mol. st. of approx. 42-
XX 45kDa with an approx. 29 AA signal sequence. The N-terminal sequence of
XX the monomeric PI is represented in each of the six repeated domains in
XX the predicted sequence of the PI precursor protein. Thus, it is likely
XX that the PI precursor protein is cleaved at six sites to produce seven
XX peptides. Six of the seven peptides, peptides 2-7 (AAR54125-R54130) would
XX be in the same mol. wt. range as the monomeric PI (about 6kDa) and would
XX have the same N-terminal sequence. Peptide 7 does not contain a consensus
XX site for trypsin or chymotrypsin. Peptide 1 (AAR54124) is small than
XX 6kDa, has a different N-terminus and was not detected in a purified
XX monomeric PI prep. It could be envisaged that peptide 1 and peptide 7
XX would form a functional proteinase inhibitor with the inhibitory site on
XX peptide 1 held in the correct conformation by disulphide bonds between
XX the two peptides. A monomer which corresp. to any of the peptides 2-7 is
XX claimed. (Updated on 25-MAR-2003 to correct PN field.)
XX CC
SQ Sequence 58 AA;

Query Match 23.7%; Score 203; DB 2; Length 58;
Best Local Similarity 63.0%; Pred. No. 8.2e-10;
Matches 34; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 53 ICTNCCSGYKGNYSAGKDLICEGESDPNPKDCTFECDTQIAYSKCPRSEK 106
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 3 ICTNCCAGKKGCKYFSDGTFVCEGESDPKNPKACPRNCDGRIAYGICPLSEK 56
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

RESULT 11
ADG69883
ID ADG69883 standard; protein; 50 AA.
XX
AC ADG69883;
XX
DT 11-MAR-2004 (first entry)
XX
DE Potato protease inhibitor II (PPI) motif structure amino acid sequence.
XX
KW transgenic plant; ferulic acid esterase; PAE; enzyme; phenolic acid;
KW plant; cell wall; improved digestibility; biomass conversion;
KW highly fermentable carbohydrate.
XX
OS Synthetic.
OS Solanum tuberosum.
XX
FN WO200268666-A1.
XX
PD 06-SEP-2002.
XX
PF 16-NOV-2001; 2001WO-US043588.
XX
PR 17-NOV-2000; 2000US-0249608P.
XX
PA (GEMV ) GENENCOR INT INC.
XX
PI Dunn-Coleman N, Langdon T, Morris P;
XX
DR WPI; 2002-698675/75.
DR N-PSDB; ADG69882.
XX

```





RESULT 15  
ADP30830  
ID ADP30830 standard; protein; 186 AA.  
XX AC  
XX ADP30830;  
XX DT  
XX 12-AUG-2004 (first entry)  
XX DE  
XX Human secreted protein SEQ ID #1597.  
XX CYTOSTATIC; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX OS  
XX Homo sapiens.  
XX PN WO2004035732-A2.  
XX PD  
XX 29-APR-2004.  
XX PF  
XX 28-AUG-2003; 2003WO-US025780.  
XX PR  
XX 29-AUG-2002; 2002US-0406576P.  
XX PR  
XX 29-AUG-2002; 2002US-0406579P.  
XX PR  
XX 29-AUG-2002; 2002US-0406585P.  
XX PR  
XX 29-AUG-2002; 2002US-0406588P.  
XX PR  
XX 29-AUG-2002; 2002US-0406608P.  
XX PR  
XX 29-AUG-2002; 2002US-0406611P.  
XX PR  
XX 29-AUG-2002; 2002US-0406612P.  
XX PR  
XX 29-AUG-2002; 2002US-0406616P.  
XX PR  
XX 29-AUG-2002; 2002US-0406640P.  
XX PR  
XX 29-AUG-2002; 2002US-0406642P.  
XX PR  
XX 29-AUG-2002; 2002US-0406646P.  
XX PR  
XX 29-AUG-2002; 2002US-0406653P.  
XX PR  
XX 29-AUG-2002; 2002US-0406655P.  
XX PR  
XX 29-AUG-2002; 2002US-0406666P.  
XX PR  
XX 17-SEP-2002; 2002US-0410946P.  
XX PR  
XX 17-SEP-2002; 2002US-0410947P.  
XX PR  
XX 17-SEP-2002; 2002US-0410948P.  
XX PR  
XX 17-SEP-2002; 2002US-0410949P.  
XX PR  
XX 17-SEP-2002; 2002US-0410953P.  
XX PR  
XX 17-SEP-2002; 2002US-0410957P.  
XX PR  
XX 17-SEP-2002; 2002US-0410958P.  
XX PR  
XX 17-SEP-2002; 2002US-0410959P.  
XX PR  
XX 17-SEP-2002; 2002US-0410960P.  
XX PR  
XX 17-SEP-2002; 2002US-0410961P.  
XX PR  
XX 17-SEP-2002; 2002US-0410962P.  
XX PR  
XX 17-SEP-2002; 2002US-0411019P.  
XX PR  
XX 17-SEP-2002; 2002US-0411022P.  
XX PR  
XX 17-SEP-2002; 2002US-0411023P.  
XX PR  
XX 17-SEP-2002; 2002US-0411024P.  
XX PR  
XX 17-SEP-2002; 2002US-0411032P.  
XX PR  
XX 17-SEP-2002; 2002US-0411035P.  
XX PR  
XX 17-SEP-2002; 2002US-0411037P.  
XX PR  
XX 17-SEP-2002; 2002US-0411041P.  
XX PR  
XX 17-SEP-2002; 2002US-0411045P.  
XX PR  
XX 17-SEP-2002; 2002US-0411046P.  
XX PR  
XX 17-SEP-2002; 2002US-0411048P.  
XX PR  
XX 17-SEP-2002; 2002US-0411052P.  
XX PR  
XX 17-SEP-2002; 2002US-0411055P.  
XX PR  
XX 17-SEP-2002; 2002US-0411073P.  
XX PR  
XX 17-SEP-2002; 2002US-0411082P.  
XX PR  
XX 17-SEP-2002; 2002US-0411101P.  
XX PR  
XX 17-SEP-2002; 2002US-0411111P.  
XX PR  
XX 18-APR-2003; 2003US-0463700P.  
XX PR  
XX 18-APR-2003; 2003US-0463708P.  
XX PR  
XX 18-APR-2003; 2003US-0463716P.  
XX PR  
XX 18-APR-2003; 2003US-0463732P.  
XX PR  
XX 02-MAY-2003; 2003US-0467199P.  
XX PR  
XX 02-MAY-2003; 2003US-0467201P.  
XX PR  
XX 02-MAY-2003; 2003US-0467203P.  
XX PR  
XX 02-MAY-2003; 2003US-0467230P.  
XX PR  
XX 19-MAY-2003; 2003US-0471306P.  
XX PR  
XX 19-MAY-2003; 2003US-0471336P.

PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kothakota S, Haishan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 2828; 428pp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytosolic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOWEB and is not in the specification.  
XX  
XX Sequence 186 AA;  
Query Match 12.2%; Score 104.5; DB 8; Length 186;  
Best Local Similarity 31.4%; Pred. No. 0.48;  
Matches 32; Conservative 4; Mismatches 43; Indels 23; Gaps 5;  
QY 29 ACTREC-GHFSYGICPRSESGSPKICTNCCSGYKGCNYYSAKG-----DLICEGSDPR 82  
Db 70 ACTACAGGGGGGCGAGGAGA----TCTTCGAGCGGCGCCCTATCTGTCCAC 125  
QY 83 NPKDCTFECDTQIAYSKCPRSEKMIKPTGTCTTCTGYGC 124  
Db 126 ATGGCGCTGCTCTCCTTC-----TTCTTCCTG--GC 154  
Search completed: January 24, 2005, 16:48:48  
Job time : 69 secs

**THIS PAGE IS BLANK**